INTRODUCTION TO LOGIC

HARRY J. GENSLER

THIRD EDITION

ROUTLEDGE

Note to E-book Users

I, the author, carefully reformatted this book to better translate into an e-book format. Sometimes this required minor changes in wording from the paper version. Since underlining often shows up poorly, underlined items (except hyperlinks) are also **bolded**. Set your e-reader to use a small or moderate size font and narrow margins; otherwise, undesirable line breaks may make proofs harder to follow. I optimized how the book displays on three Amazon devices: a Kindle Voyage, a Kindle Paperwhite, and an inexpensive 7-inch Kindle Fire; it displayed very nicely on all three (on Kindle Fire, use a WHITE background). It displayed less well, but is still usable, on an Apple iPad (here use a SEPIA background, especially for proof sections); the display may improve when the iPad Kindle app learns to handle the newer AZW3/KF8 format. I can't vouch for how this e-book displays on other devices.

Four-digit sequences like **0013** mark the beginning of a new page in the paper version of the book; so you can search for **0013** to find page 13 (or click on the tables below). If your teacher tells you to turn to page 177, for example, then you can do either of two things:

- Click SEARCH, click 123 on the keyboard to show the number keys, type **0177**, hit ENTER, and then click the search result.
- Click NOTE TO E-BOOK USERS, click the page range, and then click **177.**

Pages 1-49

10	11	12	13	14
15	16	17	18	19
20	21	22	23	24
25	26	27	28	29
30	31	32	33	34
35	36	37	38	39
40	41	42	43	44
45	46	47	48	49

Pages 50-99

50	51	52	53	54
55	56	57	58	59
60	61	62	63	64
65	66	67	68	69
70	71	72	73	74
75	76	77	78	79
80	81	82	83	84
85	86	87	88	89
90	91	92	93	94
95	96	97	98	99

Pages 100-149

100	101	102	103	104
105	106	107	108	109
110	111	112	113	114
115	116	117	118	119
120	121	122	123	124

125	126	127	128	129
130	131	132	133	134
135	136	137	138	139
140	141	142	143	144
145	146	147	148	149

Pages 150-199

150	151	152	153	154
155	156	157	158	159
160	161	162	163	164
165	166	167	168	169
170	171	172	173	174
175	176	177	178	179
180	181	182	183	184
185	186	187	188	189
190	191	192	193	194
195	196	197	198	199

Pages 200-249

200	201	202	203	204
205	206	207	208	209
210	211	212	213	214
215	216	217	218	219
220	221	222	223	224
225	226	227	228	229
230	231	232	233	234
235	236	237	238	239

240	241	242	243	244
245	246	247	248	249

Pages 250-299

250	251	252	253	254
255	256	257	258	259
260	261	262	263	264
265	266	267	268	269
270	271	272	273	274
275	276	277	278	279
280	281	282	283	284
285	286	287	288	289
290	291	292	293	294
295	296	297	298	299

Pages 300-349

300	301	302	303	304
305	306	307	308	309
310	311	312	313	314
315	316	317	318	319
320	321	322	323	324
325	326	327	328	329
330	331	332	333	334
335	336	337	338	339
340	341	342	343	344
345	346	347	348	349

Pages 350-399

350	351	352	353	354
355	356	357	358	359
360	361	362	363	364
365	366	367	368	369
370	371	372	373	374
375	376	377	378	379
380	381	382	383	384
385	386	387	388	389
390	391	392	393	394
395	396	397	398	399

Pages 400-414

400	401	402	403	404
405	406	407	408	409
410	411	412	413	414

Front pages

i	ii	iii	iv	v
vi	vii	viii	ix	x

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Introduction to Logic

Introduction to Logic is clear and concise, uses interesting examples (many philosophical in nature), and has easy-to-use proof methods. Its key features, retained in this **Third Edition**, include:

- simpler ways to test arguments, including an innovative proof method and the star test for syllogisms;
- a wide scope of materials, suiting it for introductory or intermediate courses;
- engaging examples, from philosophy and everyday life;
- useful for self-study and preparation for standardized tests, like the LSAT;
- a reasonable price (a third the cost of some competitors); and
- exercises that correspond to the free LogiCola instructional program.

This **Third Edition**:

- improves explanations, especially on areas that students find difficult;
- has a fuller explanation of traditional Copi proofs and of truth trees; and
- updates the companion LogiCola software, which now is touch friendly (for use on Windows tablets and touch monitors), installs more easily on Windows and Macintosh, and adds exercises on Copi proofs and on truth trees. You can still install LogiCola for free (from http://www.harryhiker.com/lc or http://www.routledge.com/cw/gensler).

Harry J. Gensler, S.J., is Professor of Philosophy at Loyola University Chicago. His fourteen earlier books include *Gödel's Theorem Simplified* (1984), *Formal Ethics* (1996), *Catholic Philosophy Anthology* (2005), *Historical Dictionary of Logic* (2006), *Historical Dictionary of Ethics* (2008), *Ethics: A Contemporary Introduction* (1998 & 2011), *Ethics and the Golden Rule* (2013), and *Ethics and Religion* (2016).

"Equal parts eloquent and instructive, Gensler has once again provided an invaluable resource for those looking to master the fundamental principles of logic. The Third Edition improves upon an already exceptional text by infusing the introduction of new concepts with enhanced clarity, rendering even the most challenging material a joy to teach. The updated LogicCola program is sure to become an indispensable component of my own introductory course."

Christopher Haley, Waynesburg University, USA

"This Third Edition improves on a book that was already superb. I have used Gensler's book to teach introductory courses in logic to undergraduate philosophers and linguists, and the response from the students has always been positive. They appreciate its clear explanation and the wealth of examples and practice opportunities it provides. In particular, the translation exercises help to refine logico-semantic intuitions. The supporting LogiCola software, which is feely downloadable, is a great support tool."

Mark Jary, University of Roehampton, UK

"The Third Edition is an improved version of an already excellent introduction to logic. Gensler's Reductio proof procedure enables a seamless transition from elementary propositional logic to quantification theory and more advanced modal logics. Many of the exercises involve formulations of philosophical problems. The explanations of advanced topics have been greatly improved. The upgraded LogiCola program now supports alternative proof procedures. This book is a winner!"

Michael Bradie, Bowling Green State University, USA

Introduction to Logic

Third Edition

Harry J. Gensler

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Third edition published 2017
by Routledge
711 Third Avenue, New York, NY 10017

and in the UK by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

First edition published by Routledge 2002
Second edition published by Routledge 2010

Library of Congress Cataloging in Publication Data
Gensler, Harry J., 1945 –
Introduction to logic / Harry J. Gensler. – 3rd ed.
p. cm.
Includes index.
1. Logic. I. Title.
BC71.G37 2017
160–dc22
2009039539

ISBN: 9781138910584 (hbk)
ISBN: 9781138910591 (pbk)
ISBN: 9781315693361 (ebk)

Typeset in Aldus LT Roman by the author.

Visit the companion website: http://www.harryhiker.com/lc or http://www.routledge.com/cw/gensler

Contents

Preface

This very comprehensive *Introduction to Logic* covers:

- syllogisms;
- informal aspects of reasoning (like meaning and fallacies);
- inductive reasoning;
- propositional and quantificational logic;
- modal, deontic, and belief logic;
- the formalization of an ethical view about the golden rule; and
- metalogic, history of logic, deviant logic, and philosophy of logic.

Different parts can be used in a range of logic courses, from basic introductions to graduate courses. The teachers manual and the end of Chapter 1 both talk about which chapters fit which type of course.

Earlier Routledge editions appeared in 2002 and 2011. Features included (a) clear, concise writing; (b) engaging arguments from philosophy and everyday life; (c) simpler ways to test arguments, including an innovative proof method and the syllogism star-test; (d) the widest range of materials of any logic text; (e) high suitability for self-study and preparation for tests like the LSAT; (f) a reasonable price (a third that of some competitors); and (g) the free companion LogiCola instructional program (which randomly generates problems, gives feedback on answers, provides help and explanations, and records progress). I'm happy with how earlier editions were received, often with lavish praise.

I improved this third edition in many ways. I went through the book, making explanations clearer and more concise. I especially worked on areas that students find difficult, such as (to give a few examples) why "all A is B" and "some A is not B" are contradictories (§2.4), deriving syllogistic conclusions (§2.5), the transition from inference rules to formal proofs (§§6.10–13 & 7.1), how to evaluate formulas in quantificational logic (§§8.3 & 8.5), how to translate "exactly one" and "exactly two" in identity logic (§9.1), multiple-quantifier translations and endless-loop refutations in relational logic (§§9.4–9.5), when to drop a necessary formula into the actual world in modal logic (§10.2), and how inference rules work in belief logic (§13.2). I expanded sections on traditional Copi proofs (§§7.5, 8.6, and 9.7, urged on by reviewers) and truth trees (§7.6, urged on by my friend Séamus Murphy), for teachers who might also want to teach these methods or have students learn them on their own for additional credit (as I do). "For Further Reading" now mentions further sections of the book that an advanced student might want to pursue while doing specific chapters; for example, the Basic Propositional Logic chapter goes well with sections on metalogic, deviant logic, and 000x philosophy of logic. I didn't substantially change exercise sections. Despite additions, the book is now six pages shorter.

The book now has a very nice Kindle e-book version, with real page numbers, based on a second version of the manuscript that I made with simplified formatting. And yes, you can add your own highlighting and notes.

I improved the companion LogiCola software, which runs on Windows, Macintosh, and Linux. Cloud Sync allows syncing scores between various computers. Proofs have a Training Wheels option; this gives hints about what to derive (it might bold lines 4 and 7 and ask "4 is an IF-THEN; do you have the first part true or the second part false?") – hints disappear as your score builds up. Touch features let LogiCola be done using only touch, only mouse and keyboard, or any combination of these; touch works nicely on Windows tablets or touch-screen monitors. Quantificational translations have a Hints option; this gives Loglish hints about how to translate English sentences (for "All Italians are lovers" it might say "For all x, if x is Italian then x is a lover") – hints disappear as your score builds up. There are exercises for Copi proofs and truth trees; to process scores from these, your LogiSkor program needs a version date of at least January 2016. And the Macintosh setup is easier. LogiCola (with a score-processing program, teachers manual, class slides, flash cards, and sample quizzes) can be downloaded for free from any of these Web addresses:

http://www.harryhiker.com/lc
http://www.harrycola.com/lc
http://www.routledge.com/cw/gensler

All supplementary materials are conveniently accessible from LogiCola's HELP menu; so I suggest that you just install LogiCola (teachers should check the option to install the score processor too).

I wish to thank all who have somehow contributed to this third edition. I thank Andy Beck at Routledge and his staff and reviewers, who made good suggestions. I thank my logic students, especially those whose puzzled looks pushed me to make things clearer. And I thank the many teachers, students, and self-learners who e-mailed me, often saying things like "I love the book and software, but there's one thing I have trouble with ...." If this third edition is a genuine improvement, then there are many people to thank besides me.

Long live logic!

Harry J. Gensler
Philosophy Department
Loyola University
Chicago, IL 60660 USA

http://www.harryhiker.com

0001

1 Introduction

1.1 Logic

Logic[1] is *the analysis and appraisal of arguments*. Here we'll examine reasoning on philosophical areas (like God, free will, and morality) and on other areas (like backpacking, water pollution, and football). Logic is a useful tool to clarify and evaluate reasoning, whether on deeper questions or on everyday topics.

Why study logic? First, logic builds our minds. Logic develops analytical skills essential in law, politics, journalism, education, medicine, business, science, math, computer science, and most other areas. The exercises in this book are designed to help us think more clearly (so people can better understand what we're saying) and logically (so we can better support our conclusions).

Second, logic deepens our understanding of **philosophy** – which can be defined as *reasoning about the ultimate questions of life*. Philosophers ask questions like "Why accept or reject free will?" or "Can one prove or disprove God's existence?" or "How can one justify a moral belief?" Logic gives tools to deal with such questions. If you've studied philosophy, you'll likely recognize some of the philosophical reasoning in this book. If you haven't studied philosophy, you'll find this book a good introduction to the subject. In either case, you'll get better at recognizing, understanding, and appraising philosophical reasoning.

Finally, logic can be fun. Logic will challenge your thinking in new ways and will likely fascinate you. Most people find logic enjoyable.

1.2 Valid arguments

I begin my basic logic course with a multiple-choice test. The test has ten problems; each gives information and asks what conclusion necessarily follows. The problems are fairly easy, but most students get about half wrong.[2] 0002

Here's a problem that almost everyone gets right:

[1] Key terms (like "**logic**") are introduced in bold. Learn each key term and its definition.

2 Http://www.harryhiker.com/logic.htm has my pretest in an interactive format. I suggest that you try it. I developed this test to help a psychologist friend test the idea that males are more logical than females; both groups, of course, did equally well on the problems.

If you overslept, you'll be late.
You aren't late.

Therefore

(a) You did oversleep.
(b) You didn't oversleep. ⇐ correct
(c) You're late.
(d) None of these follows.

With this next one, many wrongly pick answer "(b)":

If you overslept, you'll be late.
You didn't oversleep.

Therefore

(a) You're late.
(b) You aren't late.
(c) You did oversleep.
(d) None of these follows. ⇐ correct

Here "You aren't late" doesn't necessary follow, since you might be late for another reason; maybe your car didn't start.[1] The pretest shows that untrained logical intuitions are often unreliable. But logical intuitions can be developed; yours will likely improve as you work through this book. You'll also learn techniques for testing arguments.

In logic, an **argument** is a set of statements consisting of premises (supporting evidence) and a conclusion (based on this evidence). Arguments put reasoning into words. Here's an example ("∴" is for "therefore"):

Valid argument

If you overslept, you'll be late.
You aren't late.
∴ You didn't oversleep.

An argument is **valid** if it would be contradictory (impossible) to have the premises all true and conclusion false. "Valid" doesn't say that the premises *are* true, but only that the conclusion *follows from* them: if the premises were all true, then the conclusion would have to be true. Here we implicitly assume that there's no shift in the meaning or reference of the terms; hence we must use "overslept," "late," and "you" the same way throughout the argument.[2]

Our argument is valid because of its *logical form*: how it arranges logical

[1] These two arguments were taken from Matthew Lipman's fifth-grade logic textbook: *Harry Stottlemeier's Discovery* (Caldwell, NJ: Universal Diversified Services, 1974).

[2] It's convenient to allow arguments with zero premises; such arguments (like "∴ x = x") are valid if and only if the conclusion is a necessary truth (couldn't have been false).

notions like “if-then” and content like “You overslept.” We can display the form using words or symbols for logical notions and letters for content phrases:

If you overslept, you’ll be late.
You aren’t late.
∴ You didn’t oversleep.

If A then B **Valid**
Not-B
∴ Not-A

Our argument is valid because its *form* is correct. Replacing “A” and “B” with other content yields another valid argument of the same form: 0003

If you’re in France, you’re in Europe.
You aren’t in Europe.
∴ You aren’t in France.

If A then B **Valid**
Not-B
∴ Not-A

Logic studies forms of reasoning. The content can deal with anything – backpacking, math, cooking, physics, ethics, or whatever. When you learn logic, you’re learning tools of reasoning that can be applied to any subject.

Consider our **invalid** example:

If you overslept, you’ll be late.
You didn’t oversleep.
∴ You aren’t late.

If A then B **Invalid**
Not-A
∴ Not-B

Here the second premise denies the *first* part of the if-then; this makes it invalid. Intuitively, you might be late for some other reason – just as, in this similar argument, you might be in Europe because you’re in Italy:

If you’re in France, you’re in Europe.
You aren’t in France.
∴ You aren’t in Europe.

If A then B **Invalid**
Not-A
∴ Not-B

1.3 Sound arguments

Logicians distinguish *valid* arguments from *sound* arguments:

> An argument is **valid** if it would be contradictory to have the premises all true and conclusion false.
>
> An argument is **sound** if it's valid and every premise is true.

Calling an argument "valid" says nothing about whether its premises are true. But calling it "sound" says that it's valid (the conclusion follows from the premises) *and* has all premises true. Here's a *sound* argument:

Valid and true premises

If you're reading this, you aren't illiterate.
You're reading this.
∴ You aren't illiterate.

When we try to prove a conclusion, we try to give a *sound* argument: valid and true premises. With these two things, we have a sound argument and our conclusion has to be true.

An argument could be unsound in either of two ways: (1) it might have a false premise or (2) its conclusion might not follow from the premises: 0004

First premise false

All logicians are millionaires.
Gensler is a logician.
∴ Gensler is a millionaire.

Conclusion doesn't follow

All millionaires eat well.
Gensler eats well.
∴ Gensler is a millionaire.

When we criticize an opponent's argument, we try to show that it's *unsound*.

We try to show that one of the premises is false or that the conclusion doesn’t follow. If the argument has a false premise or is invalid, then our opponent hasn’t proved the conclusion. But the conclusion still might be true – and our opponent might later discover a better argument for it. To show a view to be false, we must do more than just refute an argument for it; we must give an argument that shows the view to be false.

Besides asking whether premises are true, we can ask how certain they are, to ourselves or to others. We’d like our premises to be certain and obvious to everyone. We usually have to settle for less; our premises are often educated guesses or personal convictions. Our arguments are only as strong as their premises. This suggests a third strategy for criticizing an argument; we could try to show that one or more of the premises are very uncertain.

Here’s another example of an argument. In fall 2008, before Barack Obama was elected US president, he was ahead in the polls. But some thought he’d be defeated by the “Bradley effect,” whereby many whites *say* they’ll vote for a black candidate but in fact don’t. Barack’s wife Michelle, in an interview with Larry King, argued that there wouldn’t be a Bradley effect:

> Barack Obama is the Democratic nominee.
> If there’s going to be a Bradley effect, then Barack wouldn’t be the nominee [because the effect would have shown up in the primaries].
> ∴ There isn’t going to be a Bradley effect.

Once she gives this argument, we can’t just say “Well, my opinion is that there *will* be a Bradley effect.” Instead, we have to respond to her reasoning. It’s clearly valid – the conclusion follows from the premises. Are the premises true? The first premise was undeniable. To dispute the second premise, we’d have to argue that the Bradley effect would appear in the final election but not in the primaries. So this argument changes the discussion. (By the way, there was no Bradley effect when Obama was elected president a month later.)

Logic, while not itself resolving substantive issues, gives us intellectual tools to reason better about such issues. It can help us to be more aware of reasoning, to express reasoning clearly, to determine whether a conclusion follows from the premises, and to focus on key premises to defend or criticize.

Logicians call statements *true* or *false* (not *valid* or *invalid*). And they call arguments *valid* or *invalid* (not *true* or *false*). While this is conventional usage, it pains a logician’s ears to hear “invalid statement” or “false argument.”[0005]

Our arguments so far have been **deductive**. With **inductive** arguments, the conclusion is only claimed to follow with probability (not with necessity):

Deductively valid

All who live in France live in Europe.
Pierre lives in France.
∴ Pierre lives in Europe.

Inductively strong

Most who live in France speak French.
Pierre lives in France.
This is all we know about the matter.
∴ Pierre speaks French (probably).

The first argument has a tight connection between premises and conclusion; it would be impossible for the premises to all be true but the conclusion false. The second has a looser premise–conclusion connection. Relative to the premises, the conclusion is only a good guess; it's likely true but could be false (perhaps Pierre is the son of the Polish ambassador and speaks no French).

1.4 The plan of this book

This book starts simply and doesn't presume any previous study of logic. Its four parts cover a range of topics, from basic to rather advanced:

- Chapters 2 to 5 cover syllogistic logic (an ancient branch of logic that focuses on "all," "no," and "some"), meaning and definitions, informal fallacies, and inductive reasoning.
- Chapters 6 to 9 cover classical symbolic logic, including propositional logic (about "if-then," "and," "or," and "not") and quantificational logic (which adds "all," "no," and "some"). Each chapter here builds on previous ones.
- Chapters 10 to 14 cover advanced symbolic systems of philosophical interest: modal logic (about "necessary" and "possible"), deontic logic (about "ought" and "permissible"), belief logic (about consistent believing and willing), and a formalized ethical theory (featuring the golden rule). Each chapter here presumes the previous symbolic ones (except that Chapter 10 depends only on 6 and 7, and Chapter 11 isn't required for 12 to 14).
- Chapters 15 to 18 cover metalogic (analyzing logical systems), history of logic, deviant logics, and philosophy of logic (further philosophical issues). These all assume Chapter 6.

Chapters 2–8 and 10 are for basic logic courses, while other chapters are more advanced. Since this book is so comprehensive, it has much more material than can be covered in one semester.

Logic requires careful reading, and sometimes rereading. Since most ideas build on previous ideas, you need to keep up with readings and problems. The companion LogiCola software (see Preface) is a great help.

2 Syllogistic Logic

Aristotle, the first logician (§16.1), invented **syllogistic logic**, which features arguments using "all," "no," and "some." This logic, which we'll take in a non-traditional way, provides a fine preliminary to modern logic (Chapters 6–14).

2.1 Easier translations

We'll now create a "syllogistic language," with rules for constructing arguments and testing validity. Here's how an English argument goes into our language:

All logicians are charming.
Gensler is a logician.
∴ Gensler is charming.

all L is C
g is L
∴ g is C

Our language uses capital letters for general categories (like "logician") and small letters for specific individuals (like "Gensler"). It uses five words: "all," "no," "some," "is," and "not." Its grammatical sentences are called **wffs**, or **well-formed formulas**. Wffs are sequences having any of these eight forms, where other capital letters and other small letters may be used instead:[1]

all A is B	x is A
no A is B	x is not A
some A is B	x is y
some A is not B	x is not y

You must use one of these exact forms (but perhaps using other capitals for "A" and "B," and other small letters for "x" and "y"). Here are examples of wffs (correct formulas) and non-wffs (misformed formulas):

[1] Pronounce "wff" as "woof" (as in "wood"). We'll take upper and lower case forms (like **A** and **a**) to be different letters, and letters with primes (like **A′** and **A′′**) to be additional letters.

Wffs: "all L is C," "no R is S," "some C is D," "g is C"

Non-wffs: "only L is C," "all R is not S," "some c is d," "G is C" 0007

Our wff rule has implications about whether to use small or capital letters:

Wffs beginning with a *word* (not a letter) use two capital letters:

Correct: "some C is D"
Incorrect: "some c is d"

Wffs beginning with a *letter* (not a word) begin with a small letter:

Correct: "g is C"
Incorrect: "G is C"

A wff beginning with a small letter could use a capital-or-small second letter (as in "a is B" or "a is b"). Which to use depends on the second term's *meaning*:

Use capital letters for **general terms**, which *describe* or put in a *category*:

B = a cute baby
C = charming
F = drives a Ford

Use capitals for "a so and so," adjectives, and verbs.

Use small letters for **singular terms**, which pick out a *specific* person or thing:

b = the world's cutest baby
t = this child
d = David

Use small letters for "the so and so," "this so and so," and proper names.

Will Gensler is a cute baby = w is B
Will Gensler is the world's cutest baby = w is b

An argument's validity can depend on whether upper or lower case is used.

Be consistent when you translate English terms into logic; use the same letter for the same idea and different letters for different ideas. It matters little which letters you use; "a cute baby" could be "B" or "C" or any other capital. I suggest that you use letters that remind you of the English terms.

Syllogistic wffs all use “is.” English sentences with a different verb should be rephrased to make “is” the main verb, and then translated. So “All dogs bark” is “all D is B” (“All dogs is [are] *barkers*”); and “Al drove the car” is “a is D” (“Al is *a person who drove the car*”).

2.1a Exercise: LogiCola A (EM & ET)[1]

Translate these English sentences into wffs.

John left the room.
j is L

1. This is a sentence.
2. This isn’t the first sentence.
3. No logical positivist believes in God.
4. The book on your desk is green. 0008
5. All dogs hate cats.
6. Kant is the greatest philosopher.
7. Ralph was born in Detroit.
8. Detroit is the birthplace of Ralph.
9. Alaska is a state.
10. Alaska is the biggest state.
11. Carol is my only sister.
12. Carol lives in Big Pine Key.
13. The idea of goodness is itself good.
14. All Michigan players are intelligent.
15. Michigan’s team is awesome.
16. Donna is Ralph’s wife.

[1] Exercise sections have a boxed sample problem that’s worked out. They also refer to LogiCola computer exercises (see Preface), which give a fun and effective way to master the material. Problems 1, 3, 5, 10, 15, and so on are worked out in the answer section at the back of the book.

2.2 The star test

Syllogisms, roughly, are arguments using syllogistic wffs. Here's an English argument and its translation into a syllogism (the Cuyahoga is a Cleveland river that used to be so polluted that it caught on fire):

No pure water is burnable.
Some Cuyahoga River water is burnable.
∴ Some Cuyahoga River water isn't pure water.

no P is B
some C is B
∴ some C is not P

More precisely, **syllogisms** are vertical sequences of one or more wffs in which each letter occurs twice and the letters "form a chain" (each wff has at least one letter in common with the wff just below it, if there is one, and the first wff has at least one letter in common with the last wff):

(If you imagine the two instances of each letter being joined, it's like a chain.)

no **P** is **B**
some **C** is **B**
∴ some **C** is not **P**

The last wff is the *conclusion*; other wffs are *premises*. Here are three more syllogisms:

a is C
b is not C
∴ a is not b

some G is F
∴ some F is G

∴ all A is A

The last example is a premise-less syllogism; it's *valid* if and only if it's impossible for the conclusion to be false.

Before doing the star test, we need to learn the technical term "distributed”:[1]

An instance of a letter is **distributed** in a wff if it occurs just after “all” or anywhere after “no” or “not.”

0009 The **<u>distributed</u>** letters below are **<u>underlined and bolded</u>**:

all **<u>A</u>** is B	x is A
no **<u>A</u>** is **<u>B</u>**	x is not **<u>A</u>**
some A is B	x is y
some A is not **<u>B</u>**	x is not **<u>y</u>**

By our definition:

- The first letter after “all” is distributed, but not the second.
- Both letters after “no” are distributed.
- Any letter after “not” is distributed.

Once you know which letters are distributed, you’re ready to learn the star test for validity. The star test is a gimmick, but a quick and effective one; for now, it’s best just to learn the test and not worry about why it works.

The **star test** for syllogisms goes as follows:

Star premise letters that are distributed and conclusion letters that aren’t distributed. Then the syllogism is **valid** if and only if every capital letter is starred *exactly* once and there is *exactly* one star on the right-hand side.

As you learn the star test, use three steps: (1) underline distributed letters, (2) star, and (3) count the stars. Here are two examples:

(1) Underline distributed letters (here only the first “A” is distributed):

all **<u>A</u>** is B
some C is A
∴ some C is B

(2) Star premise letters that are underlined and conclusion letters that aren’t

[1] §16.2 mentions the meaning of “distributed” in medieval logic. Here I suggest that you take a *distributed term* to be one that occurs just after “all” or anywhere after “no” or “not.”

underlined:

all <u>**A**</u>* is B **Valid**
some C is A
∴ some **C*** is **B***

(3) Count the stars. Here every capital letter is starred exactly once and there is exactly one star on the right-hand side. So the first argument is VALID.

(1) For our next argument, again underline distributed letters (here all the letters are distributed – since all occur after "no"):

no <u>**A**</u> is <u>**B**</u>
no <u>**C**</u> is <u>**A**</u>
∴ no <u>**C**</u> is <u>**B**</u>

(2) Star premise letters that are underlined and conclusion letters that aren't underlined:

no <u>**A**</u>* is <u>**B**</u>* **Invalid**
no <u>**C**</u>* is <u>**A**</u>*
∴ no <u>**C**</u> is <u>**B**</u>

(3) Count the stars. Here capital "A" is starred twice and there are two stars on the right-hand side. So the second argument is INVALID.

A valid syllogism must satisfy two conditions: (a) each capital letter is starred in one and only one of its instances (small letters can be starred any number of times); and (b) one and only one right-hand letter (letter after "is" or "is not") 0010 is starred. Here's an example using only small letters:

(1) Underline distributed letters (here just ones after "not" are distributed):

a is not <u>**b**</u>
∴ b is not <u>**a**</u>

(2) Star premise letters that are underlined and conclusion letters that aren't underlined:

a is not <u>**b**</u>* **Valid**
∴ **b*** is not <u>**a**</u>

(3) Count the stars. Since there are no capitals, that part is automatically satisfied; small letters can be starred any number of times. There's exactly one right-hand star. So the argument is VALID.

Here's an example without premises:

(1) Underline distributed letters:

∴ all **<u>A</u>** is A

(2) Star conclusion letters that aren't underlined:

∴ all **<u>A</u>** is **A*** **Valid**

(3) Count the stars. Each capital is starred exactly once and there's exactly one right-hand star. So the argument is VALID.

When you master this, you can skip the underlining and just star premise letters that are distributed and conclusion letters that aren't. After practice, the star test takes about five seconds to do.[1]

Logic takes "some" to mean "one or more" – and so takes this to be valid:[2]

Gensler is a logician.
Gensler is mean.
∴ Some logicians are mean.

g is L **Valid**
g is M
∴ some **L*** is **M***

Similarly, logic takes this next argument to be invalid:

Some logicians are mean.
∴ Some logicians are not mean.

some L is M **Invalid**
∴ some **L*** is not **<u>M</u>**

If *one or more* logicians are mean, it needn't be that *one or more* aren't mean; maybe *all* logicians are mean.

2.2a Exercise – No LogiCola exercise

Which of these are syllogisms?

[1] The star test is my invention. For why it works, see http://www.harryhiker.com/star.htm or my "A simplified decision procedure for categorical syllogisms," *Notre Dame Journal of Formal Logic* 14 (1973): pp. 457–66.

[2] In English, "some" can also mean "two or more," "several," "one or more but not all," "two or more but not all," or "several but not all." Only the one-or-more sense makes our argument valid.

no P is B some C is B ∴ some C is not P
This is a syllogism. (Each formula is a wff, each letter occurs twice, and the letters form a chain.)

0011

1. all C is D
∴ some C is not E

2. g is not l
∴ l is not g

3. no Y is E
all G is Y
∴ no Y is E

4. ∴ all S is S

5. k is not L
all M is L
some N is M
Z is N
∴ k is not Z

2.2b Exercise: LogiCola BH

Underline the distributed letters in the following wffs.

some R is not S
some R is not <u>**S**</u>

1. w is not s

2. some C is B

3. no R is S

4. a is C

5. all P is B

6. r is not D

7. s is w

8. some C is not P

2.2c Exercise: LogiCola B (H and S)

Valid or invalid? Use the star test.

no P is B some C is B ∴ some C is not P
no **<u>P</u>*** is **<u>B</u>*** **Valid** some C is B ∴ some **C*** is not **<u>P</u>**

1. no P is B
some C is not B
∴ some C is P

2. x is W
x is not Y
∴ some W is not Y

3. no H is B
no H is D
∴ some B is not D

4. some J is not P
all J is F
∴ some F is not P

5. ∴ g is g

6. g is not s
∴ s is not g

7. all L is M
g is not L
∴ g is not M

8. some N is T
some C is not T
∴ some N is not C

9. all C is K
s is K
∴ s is C

10. all D is A
∴ all A is D

11. s is C
s is H
∴ some C is H

12. some C is H
∴ some C is not H

13. a is b
b is c
c is d
∴ a is d

14. no A is B
some B is C
some D is not C
all D is E
∴ some E is A

2.3 English arguments

Most arguments in this book are in English. Work them out in a dual manner. First use intuition. Read the argument and ask whether it seems valid; sometimes this will be clear, sometimes not. Then symbolize the argument and do a validity 0012 test. If your intuition and the validity test agree, then you have a stronger basis for your answer. If they disagree, then something went wrong; reconsider your intuition, your translation, or how you did the validity test. This dual attack trains your logical intuitions and double-checks your results.

When you translate into logic, use the same letter for the same idea and different letters for different ideas. The same idea may be phrased in different ways;[1] often it's redundant or stilted to phrase an idea in the exact same way throughout an argument. If you have trouble remembering which letter

[1] "Express the same idea" can be tricky to apply. Consider "All Fuji apples are nutritious" and "All nutritious apples have vitamins." Use the same letter for both underlined phrases, since the first statement is equivalent to "All Fuji apples are nutritious apples."

translates which phrase, underline the phrase in the argument and write the letter above it; or write out separately which letter goes with which phrase.

Translate singular terms into small letters, and general terms into capital letters (§2.1). Capitalization can make a difference to validity. This first example uses a capital "M" (for "a man" – which could describe several people) and is invalid:

Al is a man.
My father is a man.
∴ Al is my father.

a is M **Invalid**
f is M
∴ a* is f*

This second example uses a small "m" (for "the NY mayor" – which refers to a specific person) and is valid:

Al is the NY mayor.
My father is the NY mayor.
∴ Al is my father.

a is m **Valid**
f is m
∴ a* is f*

We'll more likely catch capitalization errors if we do the problems intuitively as well as mechanically.

2.3a Exercise: LogiCola BE

Valid or invalid? First appraise intuitively. Then translate into logic and use the star test to determine validity.

No pure water is burnable. Some Cuyahoga River water is burnable. ∴ Some Cuyahoga River water isn't pure water.
no P* is B* **Valid** some C is B ∴ some C* is not P

1. All segregation laws degrade human personality.
All laws that degrade human personality are unjust.
∴ All segregation laws are unjust. [From Dr Martin Luther King.]

2. All Communists favor the poor.
All Democrats favor the poor.
∴ All Democrats are Communists. [This reasoning could persuade if expressed emotionally in a political speech. It's less likely to persuade if put into a clear premise–conclusion form.] 0013

3. All too-much-time penalties are called before play starts.
No penalty called before play starts can be refused.
∴ No too-much-time penalty can be refused.

4. No one under 18 is permitted to vote.
No faculty member is under 18.
The philosophy chairperson is a faculty member.
∴ The philosophy chairperson is permitted to vote. [Applying laws, like ones about voting, requires logical reasoning. Lawyers and judges need to be logical.]

5. All acts that maximize good consequences are right.
Some punishing of the innocent maximizes good consequences.
∴ Some punishing of the innocent is right. [This argument and the next give a mini-debate on utilitarianism. Moral philosophy would try to evaluate the premises; logic just focuses on whether the conclusion follows.]

6. No punishing of the innocent is right.
Some punishing of the innocent maximizes good consequences.
∴ Some acts that maximize good consequences aren't right.

7. All huevos revueltos are buenos para el desayuno.
All café con leche is bueno para el desayuno.
∴ All café con leche is huevos revueltos. [To test whether this argument is valid, you don't have to understand its meaning; you only have to grasp the form. In doing formal logic, you don't have to know what you're talking about; you only have to know the logical form of what you're talking about.]

8. The belief that there's a God is unnecessary to explain our experience.
All beliefs unnecessary to explain our experience ought to be rejected.
∴ The belief that there's a God ought to be rejected. [St Thomas Aquinas mentioned this argument in order to dispute the first premise.]

9. The belief in God gives practical life benefits (courage, peace, zeal, love, …).
All beliefs that give practical life benefits are pragmatically justifiable.
∴ The belief in God is pragmatically justifiable.   [From William James.]

10. All sodium salt gives a yellow flame when put into the flame of a Bunsen burner.
This material gives a yellow flame when put into the flame of a Bunsen burner.
∴ This material is sodium salt.

11. All abortions kill innocent human life.
No killing of innocent human life is right.
∴ No abortions are right.

12. All acts that maximize good consequences are right.
All socially useful abortions maximize good consequences.
∴ All socially useful abortions are right.

13. That drink is transparent.
That drink is tasteless.
All vodka is tasteless.
∴ Some vodka is transparent. 0014

14. Judy isn't the world's best cook.
The world's best cook lives in Detroit.
∴ Judy doesn't live in Detroit.

15. All men are mortal.
My mother is a man.
∴ My mother is mortal.

16. All gender-neutral terms can be applied naturally to individual women.
The term "man" can't be applied naturally to individual women. [We can't naturally say "My mother is a man"; see the previous argument.]
∴ The term "man" isn't a gender-neutral term.   [From Janice Molton.]

17. Some moral questions are controversial.
No controversial question has a correct answer.
∴ Some moral questions don't have a correct answer.

18. The idea of a perfect circle is a human concept.
The idea of a perfect circle doesn't derive from sense experience.
All ideas gained in our earthly existence derive from sense experience.
∴ Some human concepts aren't ideas gained in our earthly existence.   [This reasoning led Plato to think that the soul gained ideas in a previous existence.]

19. All beings with a right to life are capable of desiring continued existence.
All beings capable of desiring continued existence have a concept of themselves as a continuing subject of experiences.
No human fetus has a concept of itself as a continuing subject of experiences.
∴ No human fetus has a right to life. [From Michael Tooley.]

20. The bankrobber wears size-twelve hiking boots.
You wear size-twelve hiking boots.
∴ You're the bankrobber. [This is circumstantial evidence.]

21. All moral beliefs are products of culture.
No products of culture express objective truths.
∴ No moral beliefs express objective truths.

22. Some books are products of culture.
Some books express objective truths.
∴ Some products of culture express objective truths. [How can we make this valid?]

23. Dr Martin Luther King believed in objective moral truths (like "Racism is wrong").
Dr Martin Luther King disagreed with the moral beliefs of his culture.
No people who disagree with the moral beliefs of their culture are absolutizing the moral beliefs of their own culture.
∴ Some who believed in objective moral truths aren't absolutizing the moral beliefs of their own culture.

24. All claims that would still be true if no one believed them are objective truths.
"Racism is wrong" would still be true if no one believed it.
"Racism is wrong" is a moral claim.
∴ Some moral claims are objective truths. 0015

25. Some shivering people with uncovered heads have warm heads.
All shivering people with uncovered heads lose much heat through their heads.
All who lose much heat through their heads ought to put on a hat to stay warm.
∴ Some people who have warm heads ought to put on a hat to stay warm.

2.3b Mystery story exercise – No LogiCola exercise

Herman had a party at his house. Alice, Bob, Carol, David, George, and others were there; one or more of these stole money from Herman's bedroom. You have the data in the box, which may or may not give conclusive evidence about a given suspect:

1. Alice doesn't love money.
2. Bob loves money.
3. Carol knew where the money was.
4. David works for Herman.
5. David isn't the nastiest person at the party.
6. All who stole money love money.
7. All who stole money knew where the money was.
8. All who work for Herman hate Herman.
9. All who hate Herman stole money.
10. The nastiest person at the party stole money.

Did Alice steal money? If you can, prove your answer using a valid syllogism with premises from the box.

Alice *didn't* steal money:

a is not L* – #1
all S* is L – #6
∴ a* is not S

1. Did Bob steal money? If you can, prove your answer using a valid syllogism with premises from the box.

2. Did Carol steal money? If you can, prove your answer using a valid syllogism with premises from the box.

3. Did David steal money? If you can, prove your answer using a valid syllogism with premises from the box.

4. Based on our data, did more than one person steal money? Can you prove this using syllogistic logic?

5. Suppose that, from our data, we could deduce both that a person stole money and that this same person didn't steal money. What would that show?

2.4 Harder translations

Suppose we want to test this argument: 0016

Every human is mortal.
Only humans are philosophers.
∴ Every philosopher is mortal.

all H is M
all P is H

∴ all P is M

Here we need to translate "every" and "only" into our standard "all," "no," and "some." "Every" just means "all." "Only" is trickier; "Only humans are philosophers" really means "All philosophers are humans," and so it symbolizes as "all P is H" (switching the letters).

This box lists some common ways to say "all":

"all A is B" =
Every (each, any) A is B. Whoever is A is B.
A's are B's.[1] Those who are A are B. If a person is A, then he or she is B. If you're A, then you're B.
Only B's are A's. None but B's are A's.
No one is A unless he or she is B. No one is A without being B. A thing isn't A unless it's B. It's false that some A is not B.

"Only" and "none but" require switching the order of the letters:

Only dogs are collies = All collies are dogs
only D is C = all C is D

So "only" translates as "all," but with the terms reversed; "none but" works the same way. "No … unless" is tricky too, because it really means "all":

Nothing is a collie unless it's a dog = All collies are dogs
nothing is C unless it's D = all C is D

Don't reverse the letters here; only reverse with "only" and "none but."

This box lists some common ways to say "no A is B":

[1] Logicians standardly take "A's are B's" to mean "all A is B" – even though in ordinary English it also could mean "most A is B" or "some A is B."

"no A is B" =

A's aren't B's.
Every (each, any) A is non-B.
Whoever is A isn't B.
If a person is A, then he or she isn't B.
If you're A, then you aren't B.

No one that's A is B.
There isn't a single A that's B.
Not any A is B.
It's false that there's an A that's B.
It's false that some A is B.

Never use "all A is not B." Besides not being a wff, this form is ambiguous. "All
0017 cookies are not fattening" could mean "No cookies are fattening" or "Some cookies are not fattening."

These last two boxes give ways to say "some":

some A is B =

A's are sometimes B's.
One or more A's are B's.
There are A's that are B's.
It's false that no A is B.

some A is not B =

One or more A's aren't B's.
There are A's that aren't B's.
Not all A's are B's.
It's false that all A is B.

Formulas "all A is B" and "some A is not B" are *contradictories*: saying that one is false is equivalent to saying that the other is true. Here's an example:

Not all of the pills are white = Some of the pills aren't white

Similarly, "some A is B" and "no A is B" are contradictories:

It’s false that some pills are black = No pills are black

Such idiomatic sentences can be difficult to untangle. Our rules cover most cases. If you find an example that our rules don’t cover, puzzle out the meaning yourself; try substituting concrete terms, like “pills” and “white,” as above.

2.4a Exercise: LogiCola A (HM & HT)

Translate these English sentences into wffs.

Nothing is worthwhile unless it’s difficult.
all W is D

1. Only free actions can justly be punished.
2. Not all actions are determined.
3. Socially useful actions are right.
4. None but Democrats favor the poor.
5. At least some of the shirts are on sale.
6. Not all of the shirts are on sale.
7. No one is happy unless they are rich.[1]
8. Only rich people are happy.
9. Every rich person is happy.
10. Not any selfish people are happy. 0018
11. Whoever is happy is not selfish.
12. Altruistic people are happy.
13. All of the shirts (individually) cost $20.
14. All of the shirts (together) cost $20.
15. Blessed are the merciful.
16. I mean whatever I say.
17. I say whatever I mean.

[1] How would you argue against 7 to 9? Would you go to the rich part of town and find a rich person who is miserable? Or would you go to the poor area and find a poor person who is happy?

18. Whoever hikes the Appalachian Trail (AT) loves nature.
19. No person hikes the AT unless he or she likes to walk.
20. Not everyone who hikes the AT is in great shape.

2.5 Deriving conclusions

This next exercise gives you premises and has you derive a conclusion that follows validly. Do the problems in a dual manner: first try intuition, then use rules. Using intuition, read the premises slowly, say "therefore" to yourself, hold your breath, and hope that the conclusion comes. If you get a conclusion, write it down; then symbolize the argument and test for validity using the star test.

The rule approach uses four steps based on the star test:

1. Translate the premises, star, see if rules are broken.
2. Figure out the conclusion letters.
3. Figure out the conclusion form.
4. Add the conclusion, do the star test.

(1) Translate the premises into logic, star the distributed letters, and see if rules are broken. If you have two right-hand stars, or a capital letter that occurs twice without being starred exactly once, then no conclusion validly follows – so you can write "no conclusion" and stop.

(2) The conclusion letters are the two letters that occur just once in the premises. So if your premises are "x is A" and "x is B," then "A" and "B" will occur in the conclusion.

(3) Figure out the form of the conclusion:

- *If both conclusion letters are capitals*: use an "all" or "no" conclusion if every premise starts with "all" or "no"; otherwise use a "some" conclusion.
- *If at least one conclusion letter is small*: the conclusion will have a small letter, "is" or "is not," and then the other letter.
- Always derive a negative conclusion if any premise has "no" or "not."

Here are examples using "all" and "no":

- From premises "all" and "all," derive an "all" conclusion.
- From "all" and "no," derive "no." (The order of the premises doesn't matter; 0019 so from "no" and "all," also derive "no.")

Any "some" premise gives you a "some" conclusion:

- From "all" and "some" (positive), derive "some."
- From "all" and "some is not," derive "some is not."
- From "no" and "some" (positive), derive "some is not."

If the premises have a small letter but the conclusion has to have two capitals, derive "some":

- From "x is A" and "x is B," derive "some A is B."
- From "x is A" and "x is not B," derive "some A is not B."

And if at least one conclusion letter has to be small, then the conclusion will have a small letter, "is" or "is not," and then the other letter:

- From "a is b" and "b is c," derive "a is c."
- From "a is b" and "b is C," derive "a is C."
- From "a is C" and "b is not C," derive "a is not b."

Always derive a negative conclusion if any premise has "no" or "not."

(4) Add the conclusion and do the star test; if it's invalid, see if you can make it valid by reversing the letters in the conclusion (e.g., changing "all A is B" to "all B is A" – or "some A is not B" to "some B is not A" – the order matters with these two forms). Finally, put the conclusion back into English.

Suppose we want to derive a valid conclusion using all the English premises on the left. We first translate the premises into logic and star:

Some cave dwellers use fire.
All who use fire have intelligence.

---

some C is F
all F* is I

No rules are broken. "C" and "I" will occur in the conclusion. The conclusion form will be "some … is …." We find that "some C is I" follows validly, and so we can conclude "Some cave dwellers have intelligence." Equivalently, we could conclude "Some who have intelligence are cave dwellers."

Or suppose we want to derive a valid conclusion using all of these next premises. Again, we first translate the premises into logic and star:

No one held for murder is given bail.
Smith isn't held for murder.

---

no M* is B*
s is not M*

Here "M" is starred twice and there are two right-hand stars, and so rules are broken. So no conclusion follows. Do you intuitively want to conclude "Smith is given bail"? Maybe Smith is held for kidnapping and so is denied bail. 0020

Let's take yet another example:

Gensler is a logician.
Gensler is mean.

---

g is L
g is M

No rules are broken. "L" and "M" will occur in the conclusion. The conclusion form will be "some … is …." Since "some L is M" follows validly, and we can conclude "Some logicians are mean." Equivalently, we could conclude "Some who are mean are logicians."

2.5a Exercise: LogiCola BD

Derive a conclusion in English (not in wffs) that follows validly from and uses all the premises. Write "no conclusion" if no such conclusion validly follows.

No pure water is burnable. Some Cuyahoga River water is not burnable. ————————
no P* is B* some C is not B* ———————— no conclusion

Do you want to conclude "Some Cuyahoga River water is pure water"? Maybe all of the river is polluted by something that doesn't burn.

1. All human acts are determined (caused by prior events beyond our control).
No determined acts are free.

---

2. Some human acts are free.
No determined acts are free.

3. All acts where you do what you want are free.
Some acts where you do what you want are determined.

4. All men are rational animals.
No woman is a man.

5. All philosophers love wisdom.
John loves wisdom.

6. Luke was a gospel writer.
Luke was not an apostle.

0021
7. All cheap waterproof raincoats block the escape of sweat.
No raincoat that blocks the escape of sweat keeps you dry when hiking uphill.

8. All that is or could be experienced is thinkable.
All that is thinkable is expressible in judgments.
All that is expressible in judgments is expressible with subjects and predicates.
All that is expressible with subjects and predicates is about objects and properties.

9. All moral judgments influence our actions and feelings.
Nothing from reason influences our actions and feelings.

10. No feelings that diminish when we understand their origins are rational.
All culturally taught racist feelings diminish when we understand their origin.

11. I weigh 180 pounds.
My mind does not weigh 180 pounds.

12. No acts caused by hypnotic suggestion are free.
Some acts where you do what you want are caused by hypnotic suggestion.

13. All unproved beliefs ought to be rejected.
"There is a God" is an unproved belief.

14. All unproved beliefs ought to be rejected.
"All unproved beliefs ought to be rejected" is an unproved belief.

15. Jones likes raw steaks.
Jones likes champagne.

16. Some human beings seek self-destructive revenge.
No one seeking self-destructive revenge is motivated only by self-interest.
All purely selfish people are motivated only by self-interest.

17. All virtues are praised.
No emotions are praised.

0022
18. God is a perfect being.
All perfect beings are self-sufficient.
No self-sufficient being is influenced by anything outside of itself.

19. God is a perfect being.
All perfect beings know everything.
All beings that know everything are influenced by everything.

20. All basic moral norms hold for all possible rational beings as such.
No principles based on human nature hold for all possible rational beings as such.

21. All programs that discriminate simply because of race are wrong.
All racial affirmative action programs discriminate simply because of race.

22. Some racial affirmative action programs are attempts to make amends for past injustices toward a given group.
No attempts to make amends for past injustices toward a given group discriminate simply because of race. (They discriminate because of past injustices.)

23. Some actions approved by reformers are right.

Some actions approved by society aren't approved by reformers.

---

24. Some wrong actions are errors made in good faith.
No error made in good faith is blameworthy.

---

25. All moral judgments are beliefs whose correctness cannot be decided by reason.
No objective truths are beliefs whose correctness cannot be decided by reason.

Here 1–3 defend three classic views on free will: hard determinism, indeterminism, and soft determinism; 8 and 20 are from Immanuel Kant; 9 is from David Hume; 10 is from Richard Brandt; 17 and 18 are from Aristotle; and 19 is from Charles Hartshorne.

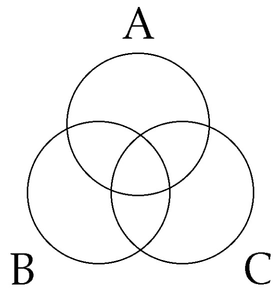
2.6 Venn diagrams

Having learned the star test, we'll now learn a second test that's more difficult but also more intuitive. **Venn diagrams** have you diagram the premises using three overlapping circles. We'll apply Venn diagrams only to *traditional syllogisms* (two-premise syllogisms with no small letters). 0023

Here's how to do the Venn-diagram test:

> Draw three overlapping circles, labeling each with one of the syllogism's letters. Then draw the premises as directed below. The syllogism is **valid** if and only if drawing the premises *necessitates* drawing the conclusion.

First, draw three overlapping circles:



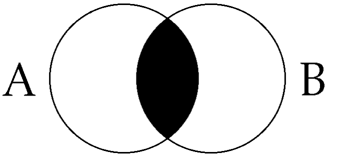
Circle A contains all A things, circle B contains all B things, and circle C contains all C things.

The central area, where all three circles overlap, contains whatever has all three features (A, B, and C). *Three middle areas* contain whatever has only two features (for example, A and B but not C). *Three outer areas* contain whatever has only one feature (for example, A but not B or C). Each of the seven areas can be empty or non-empty. We shade areas known to be empty. We put an "×" in areas known to contain at least one entity. An area without either shading or an "×" is unspecified; it could be either empty or non-empty.

Draw the premises as follows:

"no A is B"

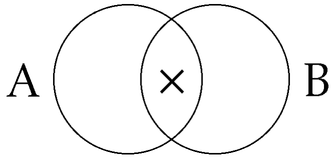
Shade wherever A and B overlap.



"No animals are beautiful" = "nothing in the animal circle is in the beautiful circle."

"some A is B"

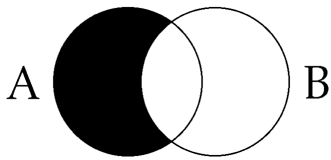
"×" an unshaded area where A and B overlap.



"Some animals are beautiful" = "something in the animal circle is in the beautiful circle."

"all A is B"

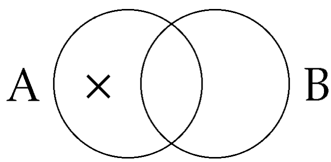
Shade areas of A that aren't in B.



"All animals are beautiful" = "everything in the animal circle is in the beautiful circle."

"some A is not B"

"×" an unshaded area in A that isn't in B.



"Some animals are not beautiful" = "something in the animal circle is outside the beautiful circle."

0024 Shading means the area is empty while "×" means it contains something.

Follow these four steps (for now you can ignore the italicized complication):

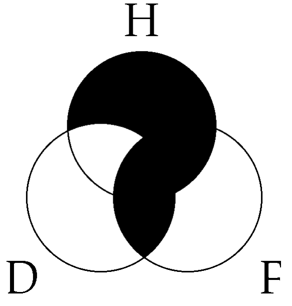
1. Draw three overlapping circles, each labeled by one of the letters.
2. First draw "all" and "no" premises by shading.
3. Then draw "some" premises by putting an "×" in some unshaded area. *(When "×"could go in either of two unshaded areas, the argument is invalid; to show this, put "×" in an area that doesn't draw the conclusion. I suggest*

you first put "×" in both areas and then erase the "×" that draws the conclusion.)

4. If you must draw the conclusion, the argument is valid; otherwise, it's invalid.

Here's a valid example:

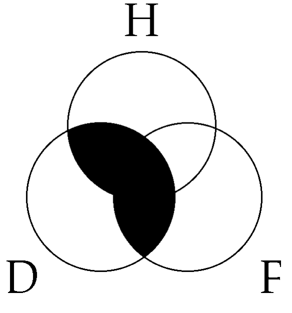
all H is D **Valid**
no F is D
∴ no H is F



We draw "all H is D" by shading areas of H that aren't in D. And we draw "no F is D" by shading where F and D overlap. Here we've automatically drawn the conclusion "no H is F" (we've shaded where H and F overlap).

So the argument is valid.
Here's an invalid example:

no H is D **Invalid**
no F is D
∴ no H is F

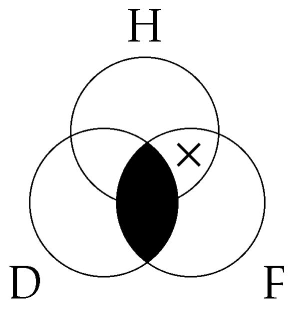


We draw "no H is D" by shading where H and D overlap. We draw "no F is D" by shading where F and D overlap. Here we haven't automatically drawn the conclusion "no H is F" (we haven't shaded all the areas where H and F overlap).

So the argument is invalid.

Here's a valid argument using "some":

no D is F **Valid**
some H is F
∴ some H is not D

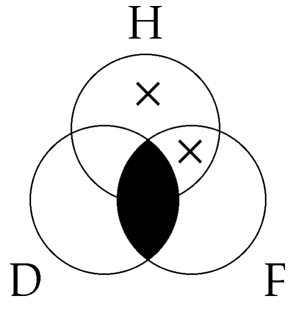


We draw "no D is F" by shading where D and F overlap. We draw "some H is F" by putting "×" in some unshaded area where H and F overlap. But then we've automatically drawn the conclusion "some H is not D" – since we've put an "×" in some area of H that's outside D.

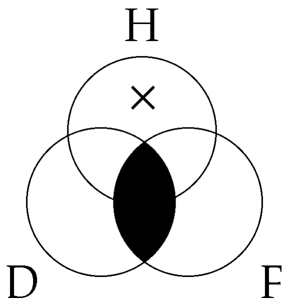
So the argument is valid. (Recall that we draw "all" and "no" first, and then "some.") 0025

I earlier warned about a complication that sometimes occurs: "*When '×' could go in either of two unshaded areas, the argument is invalid; to show this, put '×' in an area that doesn't draw the conclusion. I suggest you first put '×' in both areas and then erase the '×' that draws the conclusion.*" Here's an example:

no D is F **Invalid**
some H is not D
∴ some H is F



We draw "no D is F" by shading where D and F overlap. We draw "some H is not D" by putting "×" in *both* unshaded areas in H that are outside D (since either "×" would draw the premise).



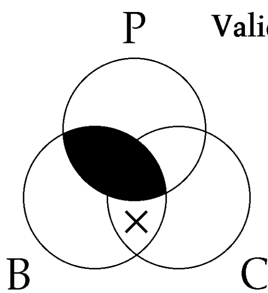
We then erase the "×" that draws the conclusion "some H is F." So then we've drawn the premises without drawing the conclusion. So it's invalid.

Since it's possible to draw the premises without drawing the conclusion, the argument is invalid. Since this case is tricky, you might reread the explanation a couple of times until it's clear in your mind.

2.6a Exercise: LogiCola BC

Test for validity using Venn diagrams.

no P is B
some C is B
∴ some C is not P



Valid

1. no B is C
all D is C
∴ no D is B

2. no Q is R
some Q is not S
∴ some S is R

3. all E is F
some G is not F
∴ some G is not E

4. all A is B
some C is B
∴ some C is A

5. all A is B
all B is C
∴ all A is C

6. all P is R
some Q is P
∴ some Q is R

7. all D is E
some D is not F
∴ some E is not F

8. all K is L
all M is L
∴ all K is M

9. no P is Q
all R is P
∴ no R is Q 0026

10. some V is W
some W is Z
∴ some V is Z

11. no G is H
some H is I
∴ some I is not G

12. all E is F
some G is not E
∴ some G is not F

2.7 Idiomatic arguments

Our arguments so far have been phrased in a clear premise–conclusion format. Real-life arguments are seldom so neat and clean. Instead we may find convoluted wording or extraneous material. Key premises may be omitted or only hinted at. And it may be hard to pick out the premises and conclusion. It often takes hard work to reconstruct a clearly stated argument from a passage.

Logicians like to put the conclusion (here italicized) last:

"Socrates is human. All humans are mortal. *So Socrates is mortal.*"

s is H
all H is M
∴ s is M

But people sometimes put the conclusion first, or in the middle:

"*Socrates must be mortal.* After all, he's human and all humans are mortal."

"Socrates is human. *So he must be mortal* – since all humans are mortal."

Here "must" and "so" indicate the conclusion (which goes last when we translate into logic). Here are some words that help us pick out premises and conclusion:

These often indicate premises:
Because, for, since, after all … I assume that, as we know … For these reasons …
These often indicate conclusions:
Hence, thus, so, therefore … It must be, it can't be … This proves (or shows) that …

When you don't have this help, ask yourself what is argued *from* (these are the premises) and what is argued *to* (this is the conclusion).

In reconstructing an argument, first pick out the conclusion. Then symbolize the premises and conclusion; this may involve untangling idioms like "Only A's are B's" (which translates as "all B is A"). If some letters occur only once, you may have to add unstated but implicit premises; using the "principle of charity," interpret unclear reasoning to give the best argument. Then test for validity.

Here's a twisted argument – and how it goes into premises and a conclusion: 0027

"You aren't allowed in here! After all, only members are allowed."

Only members are allowed in here.
∴ You aren't allowed in here.

all A is M
∴ u is not A

Since "M" and "u" occur only once, we need to add an implicit premise linking these to produce a syllogism. We add a plausible premise and test for validity:

You aren't a member. (implicit)
Only members are allowed in here.
∴ You aren't allowed in here.

u is not M* **Valid**
all A* is M
∴ u* is not A

2.7a Exercise: LogiCola B (F & I)

First appraise intuitively. Then pick out the conclusion, translate into logic (using correct wffs and syllogisms), and determine validity using the star test. Supply implicit premises where needed; when two letters occur only once but stand for different ideas, we often need an implicit premise that connects the two.

Whatever is good in itself ought to be desired. But whatever ought to be desired is capable of being desired. So only pleasure is good in itself, since only pleasure is capable of being desired.
all G* is O **Valid** all O* is C all C* is P ∴ all G is P* The conclusion is "*Only pleasure is good in itself*": "all G is P."

1. Racial segregation in schools generates severe feelings of inferiority among black students. Whatever generates such feelings treats students unfairly on the basis of race. Anything that treats students unfairly on the basis of race violates the 14th Amendment. Whatever violates the 14th Amendment is unconstitutional. Thus racial segregation in schools is unconstitutional. [This was the reasoning behind the 1954 *Brown vs. Topeka Board of Education* Supreme Court decision.]

2. You couldn't have studied! The evidence for this is that you got an F– on the test.

3. God can't condemn agnostics for non-belief. For God is all-good, anyone who is all-good respects intellectual honesty, and no one who does this condemns agnostics for non-belief.

4. Only what is under a person's control is subject to praise or blame. Thus the consequences of an action aren't subject to praise or blame, since not all the consequences of an action are under a person's control.

5. No synthetic garment absorbs moisture. So no synthetic garment should be worn next to the skin while skiing.

6. Not all human concepts can be derived from sense experience. My reason for saying this is that the idea of "self-contradictory" is a human concept but isn't derived from sense experience. 0028

7. Analyses of humans in purely physical-chemical terms are neutral about whether we have inner consciousness. So, contrary to Hobbes, we must conclude that no analysis of humans in purely physical-chemical terms fully explains our mental activities. Clearly, explanations that are neutral about whether we have inner consciousness don't fully explain our mental activities.

8. Only what is based on sense experience is knowledge about the world. It follows that no mathematical knowledge is knowledge about the world.

9. Not all the transistors in your radio can be silicon. After all, every transistor that works well at high temperatures is silicon and yet not all the transistors in your radio work well at high temperatures.

10. Moral principles aren't part of philosophy. This follows from these considerations: Only objective truths are part of philosophy. Nothing is an objective truth unless it's experimentally testable. Finally, of course, moral principles aren't experimentally testable. [From the logical positivist A. J. Ayer.]

11. At least some women are fathers. This follows from these facts: (1) Jones is a father, (2) Jones had a sex change to female, and (3) whoever had a sex change to female is (now) a woman.

12. Only language users employ generalizations. Not a single animal uses language. At least some animals reason. So not all reasoners employ generalizations. [From John Stuart Mill.]

13. Only pure studies in form have true artistic worth. This proves that a thing doesn't have true artistic worth unless it's abstract, for it's false that there's something that's abstract but that isn't a pure study in form.

14. Anything that relieves pressure on my blisters while I hike would allow me to finish my PCT (Pacific Crest Trail) hike from Mexico to Canada. Any insole with holes cut out for blisters would relieve pressure on my blisters while I hike. I conclude that any insole with holes cut out for blisters would allow me to finish my PCT hike from Mexico to Canada. [So I reasoned – and it worked.]

15. We know (from observing the earth's shadow on the moon during a lunar eclipse) that the earth casts a curved shadow. But spheres cast curved shadows. These two facts prove that the earth is a sphere.

16. Whatever is known is true, and whatever is true corresponds to the facts. We may conclude that no belief about the future is known.

17. No adequate ethical theory is based on sense experience, because any adequate ethical theory provides necessary and universal principles, and nothing based on sense experience provides such principles. [From Immanuel Kant.]

18. At least some active people are hypothermia victims. Active people don't shiver. It follows that not all hypothermia victims shiver. [From a ski magazine.]

19. Iron objects conduct electricity. We know this from what we learned last week – namely, that iron objects are metallic and that nothing conducts electricity unless it's metallic.

20. Only things true by linguistic convention are necessary truths. This shows that "God exists" can't be a necessary truth. After all, existence claims aren't true by linguistic convention.

21. No bundle of perceptions eats food. Hume eats food, and Hume is a human person. From this it follows (contrary to David Hume's theory) that no human person is a bundle of perceptions. 0029

22. Any events we could experience as empirically real (as opposed to dreams or hallucinations) could fit coherently into our experience. So an uncaused event couldn't be experienced as empirically real. I assume that it's false that some uncaused event could fit coherently into our experience. [From Immanuel Kant.]

23. I think I'm seeing a chair. But some people who think they're seeing a chair are deceived by their senses. And surely people deceived by their senses don't really know that they're seeing an actual chair. So I don't really know that I'm seeing an actual chair.

24. No material objects can exist unperceived. I say this for three reasons: (1) Material objects can be perceived. (2) Only sensations can be perceived. Finally, (3) no sensation can exist unperceived. [Bertrand Russell criticized this argument for an idealist metaphysics.]

25. Only those who can feel pleasure or pain deserve moral consideration. Not all plants can feel pleasure or pain. So not all plants deserve moral consideration.

26. True principles don't have false consequences. There are plausible principles with false consequences. Hence not all true principles are plausible.

27. Only what divides into parts can die. Everything that's material divides into parts. No human soul is material. This shows that no human soul can die.

2.8 The Aristotelian view

Historically, “Aristotelian” and “modern” logicians disagree about the validity of some syllogism forms. They disagree because of differing policies about allowing *empty terms* (general terms that don’t refer to any existing beings).

Compare these two arguments (unicorns don’t really exist, even though some myths speak of such one-horned horse-like animals):

All cats are animals.
∴ Some animals are cats.

All unicorns are animals.
∴ Some animals are unicorns.

The first seems valid while the second seems invalid. Yet both have the same form – one that tests out as “invalid” using our star test:

all C* is A   Invalid
∴ some A* is C*

When we read the first argument, we tend to presuppose that there’s at least one cat. Given this as an assumed additional premise, it follows validly that some animals are cats. When we read the second argument, we don’t assume that there’s at least one unicorn. Without this additional assumption, it doesn’t follow that some animals are unicorns.

So “all C is A ∴ some A is C” is *valid* if we assume as a further premise that there are C’s; it’s *invalid* if we don’t assume this. The *Aristotelian view*, which assumes that each general term in a syllogism refers to at least one existing
0030 being, calls the argument “valid.” The *modern view*, which allows empty terms like “unicorn” that don’t refer to existing beings, calls the argument “invalid.”

I prefer the modern view, since we often don’t presuppose that our general terms refer to existing entities. If your essay argues that angels don’t exist, your use of “angel” doesn’t presuppose that there are angels. If you tell your class “All with straight-100s may skip the final exam,” you don’t assume that anyone will get straight-100s. On the other hand, we sometimes *can* presuppose that our general terms all refer; then the Aristotelian test makes sense.

Suppose we have an argument with true premises that’s valid on the Aristotelian view but invalid on the modern view. We should draw the conclusion if we know that each general term in the premises refers to at least one existing being; otherwise, we shouldn’t. Consider this pair of arguments with the same form (a form that’s valid on the Aristotelian view but invalid on the modern view):

All cats are mammals.
All cats are furry.
∴ Some mammals are furry.

All square circles are squares.
All square circles are circles.
∴ Some squares are circles.

The first inference is sensible, because there are cats. The second inference isn't sensible, because there are no square circles.

Some logic books use the Aristotelian view, but most use the modern view. It makes a difference to very few cases; all the syllogisms in this chapter prior to this section test out the same on either view.

To adapt the star test to the Aristotelian view, word it so that each capital letter must be starred *at least once* (instead of "exactly once"). To adapt Venn diagrams to the Aristotelian view, add this rule: "If you have a circle with only one unshaded area, put an '×' in this area"; this assumes that the circle isn't empty.

3 Meaning and Definitions

Since we need to know what premises *mean* before we can appraise their truth, language is important for appraising arguments. Imagine someone giving this argument, which is deductively valid but has obscure premises:

If there's a cosmic force, then there's a God.
There's a cosmic force.
∴ There's a God.

What does the person mean by "cosmic force"? We can't intelligently agree or disagree with premises if we don't understand what they mean.

In this chapter, after looking at general uses of language, we'll examine definitions and other ways to clarify meaning. Then we'll talk about making distinctions and detecting unclarities. Finally, we'll consider the distinction between analytic and synthetic statements, and the related distinction between knowledge based on reason and knowledge based on experience. The goal of this study of language is to enhance our ability to analyze and appraise arguments.

3.1 Uses of language

The four grammatical sentence types broadly reflect four uses of language:

Declarative (make assertions): "Michigan beat Ohio State."
Interrogative (ask questions): "Did Michigan win?"
Imperative (tell what to do): "Beat Ohio State."
Exclamatory (express feelings): "Hurrah for Michigan!"

Sentences can do various jobs at the same time. While making assertions, we can also ask questions, tell what to do, or express feelings:

"I wonder whether Michigan won." (This can ask a question.)
"I want you to throw the ball." (This can tell what to do.)
"Michigan won!" (This can express feelings of joy.)

Arguments too can exemplify different uses of language. Suppose someone

0032 argues this way about the Cleveland river that used to catch on fire: “You can see that the Cuyahoga River is polluted from the fact that it can burn!” We could make this into an explicit argument:

No pure water is burnable.
Some Cuyahoga River water is burnable.
∴ Some Cuyahoga River water isn’t pure water.

One who argues this way also can (perhaps implicitly) be raising a question, directing people to do something, or expressing feelings:

“What can we do to clean up this polluted river?”
“Let’s all resolve to take action on this problem.”
“How disgusting is this polluted river!”

Arguments have a wider human context and purpose. We should remember this when we study detached specimens of argumentation.

When we do logic, our focus narrows and we concentrate on assertions and reasoning. For this purpose, detached specimens are better. Expressing an argument in a clear, direct, emotionless way can make it easier to appraise the truth of the premises and the validity of the inference.

It’s good to avoid emotional language when we reason. Of course, there’s nothing wrong with feelings or emotional language. Reason and feeling are both important parts of life; but we often need to focus on one or the other for a given purpose. At times, expressing feelings is the important thing and argumentation only gets in the way. At other times, we need to reason things out in a cool-headed manner.

Emotional language can discourage clear reasoning. When reasoning about abortion, for example, it’s wise to avoid slanted phrases like “the atrocious crime of abortion” or “Neanderthals who oppose the rights of women.” Bertrand Russell gave this example of how we slant language: “I am *firm*; you are *obstinate*; he is *pig-headed*.” Slanted phrases can mislead us into thinking we’ve defended our view by an argument (premises and conclusion) when in fact we’ve only expressed feelings. Careful thinkers try to avoid emotional terms when constructing arguments.

In the rest of this chapter, we’ll explore aspects of the “making assertions” side of language that relate closely to analyzing and appraising arguments.

3.1a Exercise

For each word or phrase, say whether it has a positive, negative, or neutral emotional tone. Then find another word or phrase with more or less the same assertive meaning but a different emotional tone. 0033

old maid
This has a negative tone. A more neutral phrase is “elderly woman who has never married.”

(The term “old maid” suggests that a woman’s goal in life is to get married and that an older woman who has never married is unfortunate. I can’t think of a corresponding negative term for an older man who never married. A word or phrase sometimes suggests a whole attitude toward life, and often an unexamined attitude.)

1. a cop
2. filthy rich
3. heroic
4. an extremist
5. an elderly gentleman
6. a bastard
7. baloney
8. a backward country
9. authoritarian
10. a do-gooder
11. a hair-splitter
12. an egghead
13. a bizarre idea
14. a kid
15. booze
16. a gay
17. abnormal
18. bureaucracy
19. abandoning me
20. babbling
21. brazen
22. an old broad
23. old moneybags
24. a busybody
25. a bribe

26. old-fashioned

27. brave

28. garbage

29. a cagey person

30. a whore

3.2 Lexical definitions

We noted earlier that the phrase “cosmic force” in this example is obscure:

If there’s a cosmic force, then there’s a God.
There’s a cosmic force.
∴ There’s a God.

Unless the speaker tells us what is meant by “cosmic force,” we won’t be able to understand what’s said or tell whether it’s true. But how can the speaker explain what he or she means by “cosmic force”? Or, more generally, how can we explain the meaning of a word or phrase?

Definitions are an important way to explain meaning. A **definition** is a rule of paraphrase intended to explain meaning. More precisely, a definition of a word or phrase is a rule saying how to eliminate this word or phrase in a sentence and produce a second sentence that means the same thing, the purpose of this being to explain or clarify the meaning of the word or phrase.

Suppose the person with the cosmic-force argument defines “cosmic force” as “force, in the sense used in physics, whose influence covers the entire universe.” This makes the first premise doubtful, since then it only means “If there’s a force [e.g., gravity], in the sense used in physics, whose influence covers the entire universe, then there’s a God.” Why think that this premise is true?

So a definition is a rule of paraphrase intended to explain meaning. Definitions may be **lexical** (explaining current usage) or **stipulative** (specifying your own usage). Here’s a correct lexical definition: 0034

“Bachelor” means “unmarried man.”

This claims that we can interchange “bachelor” and “unmarried man” in a sentence; the resulting sentence will mean the same as the original, according to current usage. This leads to the interchange test for lexical definitions:

> **Interchange test**: To test a lexical definition claiming that A means B, try switching A and B in a variety of sentences. If some resulting pair of sentences doesn't mean the same thing, then the definition is incorrect.

According to our definition of "bachelor" as "unmarried man," for example, these two sentences would mean the same thing:

"Al is a *bachelor*"
"Al is an *unmarried man*"

These *do* seem to mean the same thing. To refute the definition, we'd have to find two sentences that are alike, except that "bachelor" and "unmarried man" are interchanged, and that don't mean the same thing.

Here's an incorrect lexical definition:

"Bachelor" means "happy man."

This leads to incorrect paraphrases. If the definition were correct, then these two sentences would mean the same thing:

"Al is a *bachelor*"
"Al is a *happy man*"

But they don't mean the same thing, since we could have one true but not the other. So the definition is wrong.

The interchange test is subject to at least two restrictions. First, definitions are often intended to cover just one sense of a word that has various meanings; we should then use the interchange test only on sentences using the intended sense. Thus it wouldn't be a good objection to our definition of "bachelor" as "unmarried man" to claim that these two sentences don't mean the same:

"I have a *bachelor* of arts degree"
"I have an *unmarried man* of arts degree"

The first sentence uses "bachelor" in a sense the definition doesn't try to cover.

Second, we shouldn't use the test on sentences where the word appears in quotes. Consider this pair of sentences: 0035

"'*Bachelor*' has eight letters"
"'*Unmarried man*' has eight letters"

The two don't mean the same thing, since the first is true and the second false. But this doesn't show that our definition is wrong.

Lexical definitions are important in philosophy. Many philosophers, from Socrates to the present, have sought correct lexical definitions for some of the central concepts of human existence. They've tried to define concepts such as knowledge, truth, virtue, goodness, and justice. Such definitions are important for understanding and applying the concepts. Defining "good" as "what society approves of" would lead us to base our ethical beliefs on what's socially approved. We'd reject this method if we defined "good" as "what I like" or "what God desires," or if we regarded "good" as indefinable.

We can evaluate philosophical lexical definitions using the interchange test; Socrates was adept at this. Consider cultural relativism's definition of "good":

"X is good" means "X is approved by my society."

To evaluate this, we'd try switching "good" and "approved by my society" in a sentence to get a second sentence. Here's such a pair of sentences:

"Slavery is *good*"
"Slavery is *approved by my society*"

Then we'd see if the two sentences mean the same thing. Here they clearly don't, since it's consistent to affirm one but deny the other. Those who disagree with the norms of their society often say things like "Slavery is approved by my society, but it's not good." Given this, we can argue against cultural relativism's definition of "good" as follows:

If cultural relativism's definition is correct, then these two sentences mean the same thing.
They don't mean the same thing.
∴ Cultural relativism's definition isn't correct.

To counter this, the cultural relativist would have to claim that the sentences *do* mean the same thing. But this claim is implausible.

Here are five rules for good lexical definitions:

1. A good lexical definition is neither too broad nor too narrow.

Defining “bachelor” as “man” is too broad, since some men aren’t bachelors. And defining “bachelor” as “unmarried male astronaut” is too narrow, since some bachelors aren’t astronauts. 0036

2. A good lexical definition avoids circularity and poorly understood terms.

Defining “true” as “known to be true” is circular, since it defines “true” using “true.” And defining “good” as “having positive aretaic value” uses poorly understood terms, since “aretaic” is less clear than “good.”

3. A good lexical definition matches in vagueness the term defined.

Defining “bachelor” as “unmarried male over 18 years old” is overly precise. “Bachelor” is vague, since the exact age that the term begins to apply is unclear on semantic grounds; so “over 18” is too precise to define “bachelor.” “Man” or “adult” are better, since these match “bachelor” fairly well in vagueness.

4. A good lexical definition matches, as far as possible, the emotional tone (positive, negative, or neutral) of the term defined.

It won’t do to define “bachelor” as “*fortunate* man who never married” or “*unfortunate* man who never married.” These have positive and negative emotional slants; the original term “bachelor” is fairly neutral.

5. A good lexical definition includes only properties essential to the term.

Suppose all bachelors live on the planet earth. Even so, living on planet earth isn’t a property *essential* to the term “bachelor,” since we could imagine a bachelor who lives on the moon. So it’s wrong to include “living on planet earth” in the definition of “bachelor.”

3.2a Exercise: LogiCola Q

Give objections to these proposed lexical definitions.

“Game” means “anything that involves competition between two parties, and winning and losing.”
By this definition, solitaire isn’t a game, but a military battle is. This goes against the normal usage of the word “game.”

1. “Lie” means “false statement.”

2. “Adolescent” means “person between 9 and 19 years old.”

3. “God” means “object of ultimate concern.”

4. “Metaphysics” means “any sleep-inducing subject.”

5. “Good” means “of positive value.”

6. “Human being” means “featherless biped.”

7. “I know that P” means “I believe that P.”

8. “I know that P” means “I believe that P, and P is true.”

9. “Chair” means “what you sit on.” 0037

10. “True” means “believed.”

11. “True” means “proved to be true.”

12. “Valid argument” means “argument that persuades.”

13. “Murder” means “killing.”

14. “Morally wrong” means “against the law.”

15. “Philosopher” means “someone who has a degree in philosophy” and “philosophy” means “study of the great philosophers.”

3.2b Exercise

Cultural relativism (CR) claims that “good” (in its ordinary usage) means “socially approved” or “approved by the majority (of the society in question).” What does this definition entail about the statements below? If this definition were correct, then would each of the following be true (1), false (0), or undecided by such considerations (?)?

If torturing people for religious beliefs is socially approved in country X, then it’s good in country X.
1 (for “true”). On cultural relativism, the statement means this (and would be true): “If torturing people for religious beliefs is socially approved in country X, then it’s socially approved in country X.”

1. Conclusions about what is good are deducible from sociological data (based, for example, on opinion surveys) describing one’s society and what it approves.

2. If I say “Infanticide isn’t good” but an ancient Roman says “Infanticide is good,” then one or the other of us must be mistaken.

3. The norms set up by my society about what is good couldn’t be mistaken.

4. Judgments about what is good aren’t true or false.

5. It’s good to respect the values of other societies.

6. If our society were to favor intolerance, then intolerance would be good.

7. Representational democracy will work anywhere.

8. From an analysis of how people use the word “good,” it can be proved that whatever is socially approved must be good.

9. Different cultures accept different moral beliefs.

10. “The majority favors this” logically entails “This is good.”

11. If the majority favors war (sexual stereotypes, conservative politics, abortion, and so on), then this has to be good.

12. “Do good” means “Do what the majority favors.”

13. Doing something because it’s good isn’t the same as doing it because the majority favors it.

14. People who said “Racism is favored by the majority but it’s not good” were contradicting themselves.

15. Something that’s bad might nevertheless be socially approved (because society may be misinformed or irrational in its evaluations).

16. The majority knows what it favors.

17. If Nazism became widespread and genocide came to be what most people favored, then genocide would have to be good.

18. It’s not necessarily good for me to do what society favors. 0038

19. Suppose a survey showed that 90 percent of the population disapprove of people always following social approval. Then it follows that it’s bad to always follow social approval – in other words, it’s bad to always follow what is good.

20. Suppose your fellow Americans as a group and your fellow Anglicans as a group disapprove of racism, whereas your fellow workers and your social group (friends and relatives) approve of racism. Then racism is bad.

3.3 Stipulative definitions

A **stipulative definition** specifies how you’re going to use a term. Since your usage may be a new one, it’s unfair to criticize a stipulative definition for clashing with conventional usage. Stipulative definitions should be judged, not as correct or incorrect, but rather as useful or useless.

This book has many stipulative definitions. I continually define terms like “logic,” “argument,” “valid,” “wff,” and so forth. These definitions specify the meaning I’m going to use for the terms, which sometimes is close to their standard meaning. The definitions create a technical vocabulary.

A **clarifying definition** is one that stipulates a clearer meaning for a vague term. For example, a scientist might stipulate a technical sense of “pure water” in terms of bacteria level; this technical sense, while related to the normal one, is more scientifically precise. Likewise, courts might stipulate a more precise definition of “death” to resolve certain legal disputes; the definition might be chosen on moral and legal grounds to clarify the law.

Philosophers often use stipulative definitions. They might say: “Here I’ll use ‘rational’ to mean ‘always adopting the means believed necessary to achieve one’s goals.’” This signals that the author will use “rational” to abbreviate a certain longer phrase; it doesn’t claim that this exactly matches the term’s ordinary meaning. Others may use “rational” in different senses, such as “logically consistent,” “emotionless,” or “forming beliefs solely by the methods of science.” These thinkers needn’t be disagreeing; they may just specify their technical vocabulary differently. We could use subscripts for different senses; “rational$_1$” might mean “logically consistent,” and “rational$_2$” might mean “emotionless.” Don’t be misled into thinking that, because being rational in one sense is desirable, therefore being rational in another sense must also be desirable.

Stipulative definitions, while they needn’t accord with current usage, should:

- use clear terms that the parties involved will understand,
- avoid circularity,
- let us paraphrase out the defined term,
- accord with how the person giving it will use the term, and
- aid our understanding and discussion of the subject matter.

A stipulative definition is a device for abbreviating language. Our Chapter 1
0039 starts with a stipulative definition: “Logic is the analysis and appraisal of arguments.” This definition lets us use the one word “logic” in place of the six words “the analysis and appraisal of arguments.” Without the definition, our explanations would be wordier and harder to grasp; so the definition is useful.

Stipulative definitions should promote understanding. It’s seldom useful to stipulate that a well-established term will be used in a radical new sense (for example, that “biology” will be used to mean “the study of earthquakes”); this would create confusion. And it’s seldom useful to multiply stipulative definitions for terms that we’ll seldom use. But when we keep repeating a cumbersome phrase over and over, a stipulative definition can be helpful. Suppose your essay keeps repeating the phrase “action that satisfies criteria 1, 2, and 3 of the previous section”; your essay may be easier to follow if some short term were stipulated to mean the same as this longer phrase.

Some of our definitions seem to violate the “avoid circularity” norm. Section 6.1 defines “wffs” as sequences constructable using these rules:

1. Any capital letter is a wff.
2. The result of prefixing any wff with "∼" is a wff.
3. The result of joining any two wffs by "•" or "∨" or "⊃" or "≡" and enclosing the result in parentheses is a wff.

Clauses 2 and 3 define "wff" in terms of "wff." And the definition doesn't seem to let us paraphrase out the term "wff"; we don't seem able to take a sentence using "wff" and say the same thing without "wff."

Actually, our definition is perfectly fine. We can rephrase it in the following way to avoid circularity and show how to paraphrase out the term "wff":

"Wff" means "member of every set S of strings that satisfies these conditions: (1) Every capital letter is a member of set S; (2) the result of prefixing any member of set S with '∼' is a member of set S; and (3) the result of joining any two members of set S by '•' or '∨' or '⊃' or '≡' and enclosing the result in parentheses is a member of set S."

Our definition of "wff" is a **recursive definition** – one that first specifies some things that the term applies to and then specifies that if the term applies to certain things, then it also applies to certain other things. Here's a recursive definition of "ancestor of mine" – followed by an equivalent non-recursive definition:

1. My father and mother are ancestors of mine.
2. Any father or mother of an ancestor of mine is an ancestor of mine.

"Ancestor of mine" means "member of every set S that satisfies these conditions: (1) my father and mother are members of S; and (2) every father or mother of a member of S is a member of S." 0040

3.4 Explaining meaning

If we avoid circular definitions, we can't define all our terms; instead, we must leave some terms undefined. But how can we explain such *undefined* terms? One way is by examples.

To teach "red" to someone who understands no language that we speak, we could point to red objects and say "Red!" We'd want to point to different kinds of red object; if we pointed only to red shirts, the person might think that "red" meant "shirt." If the person understands "not," we also could point to non-red objects and say "Not red!" The person, unless color-blind, soon will catch our meaning and be able to point to red objects and say "Red!" This is a basic, primitive way to teach language. It explains a word, not by using other words, but by relating a word to concrete experiences.

We sometimes point to examples through words. We might explain "plaid" to a child by saying "It's a color pattern like that of your brother's shirt." We might explain "love" through examples: "Love is getting up to cook a sick person's breakfast instead of staying in bed, encouraging someone instead of complaining, and listening to other people instead of telling them how great you are." It's often helpful to combine a definition with examples, so the two reinforce each other; so Chapter 1 defined "argument" and then gave examples.

In abstract discussions, people sometimes use words so differently that they communicate poorly and almost seem to speak different languages. Asking for definitions may then lead to the frustration of having one term you don't understand being defined using other terms you don't understand. It may be more helpful to ask for examples: "Give me examples of an *analytic statement* (or of a *deconstruction*)." Asking for examples can bring a bewilderingly abstract discussion back down to earth and mutual understanding.

Logical positivists and pragmatists gave other ways to clarify statements. Positivists proposed that we explain a statement's meaning by specifying which experiences would show the statement to be true or to be false. Such operational definitions connect meaning to an experimental test:

- To say that rock A is "harder than" rock B means that A would scratch B but B wouldn't scratch A.
- To say that this string is "1 meter long" means that, if you stretch it over the standard meter stick, then the ends of both will coincide.
- To say that this person "has an IQ of 100" means that the person would get an average score on a standard IQ test.

Such definitions are important in science.

Logical positivists like A. J. Ayer appealed to the *verifiability criterion of meaning* as the cornerstone of their philosophy. We can formulate their principle (to be applied only to statements not true-by-definition, see §3.6) as follows: 0041

Logical positivism (LP)

To help us find a statement's meaning, ask "How could the truth or falsity of the statement in principle be discovered by conceivable observable tests?"

If there's no way to test a statement, then it has no meaning (it makes no assertion that could be true or false). If tests are given, they specify the meaning.

There are problems with taking LP to be literally true. LP says *any untestable statement is without meaning*. But LP itself is untestable. Hence LP is without meaning on its own terms; it's self-refuting. For this reason and others, few hold this view anymore, even though it was popular decades ago.

Still, the LP way to clarify statements can sometimes be useful. Consider this claim of Thales, the ancient Greek alleged to be the first philosopher: "Water is the primal stuff of reality." The meaning here is unclear. We might ask Thales for a definition of "primal stuff"; this would clarify the claim. Or we might follow LP and ask, "How could we test whether your claim is correct?" Suppose Thales says the following, thus giving an operational definition:

> Try giving living things no water. If they die, then this proves my claim. If they live, then this refutes my claim.

We'd then understand Thales to be claiming that water is needed for life. Or suppose Thales replies this way:

> Let scientists work on the task of transforming each kind of matter (gold, rock, air, and so on) into water, and water back into each kind of matter. If they eventually succeed, then that proves my claim.

Again, this would help us understand the claim. But suppose Thales says "No conceivable experimental test could show my claim to be true or show it to be false." The positivists would immediately conclude that Thales's claim is meaningless – that it makes no factual assertion that could be true or false. We non-positivists needn't draw this conclusion so quickly; but we may remain suspicious of Thales's claim and wonder what he's getting at.

LP demands that a statement *in principle* be able to be tested. Consider "There are mountains on the other side of the moon." When the positivists wrote, rockets were less advanced and the statement couldn't be tested. But that didn't matter to its meaningfulness, since we could describe what a test would be like. That this claim was *testable in principle* was enough to make it meaningful.

LP hides an ambiguity when it speaks of "conceivable observable tests." Observable by whom? Is it enough that one person can make the observation? Or does it have to be publicly observable? Is a statement about my present feelings meaningful if I alone can observe whether it's true? Historically, most positivists demanded that a statement be *publicly verifiable*. But the weaker version 0042 of the theory that allows *verification by one person* seems better. After all, a statement about my present feelings makes sense, but only I can verify it.

William James suggested a related way to clarify statements. His "Pragmatism" essay suggests that we determine the meaning, or "cash value," of a statement by relating it to practical consequences. James's view is broader and more tolerant than that of the positivists. We can formulate his pragmatism principle as follows (again, it's to be applied only to statements not true-by-definition):

Pragmatism (PR)

To help us find a statement's meaning, ask "What conceivable practical differences to someone could the truth or falsity of the statement make?" Here "practical differences to someone" covers what *experiences* one would have or what *choices* one ought to make.

If the truth or falsity of a statement could make no practical difference to anyone, then it has no meaning (it makes no assertion that could be true or false). If practical differences are given, they specify the meaning.

I'm inclined to think that something close to PR is literally true. But here I'll just stress that PR can be useful in clarifying meaning.

PR often applies much like the weaker version of LP that allows verification by one person. LP focuses on what we could experience if the statement were true or false, while PR includes such experiences under practical differences.

PR also includes under "practical differences" what choices one ought to make. This makes PR broader than LP, since what makes a difference to choices needn't be testable by observation. Hedonism claims "Only pleasure is worth striving for." LP asks "How could the truth or falsity of hedonism in principle be discovered by conceivable observable tests?" Perhaps it can't; then LP would see hedonism as cognitively meaningless. PR asks "What conceivable practical differences to someone could the truth or falsity of hedonism make?" Here, "practical differences" include what *choices* one ought to make. The truth of hedonism could make many differences about choices; if hedonism is true, for example, then we should pursue knowledge not for its own sake but only insofar as it promotes pleasure. Ethical claims like hedonism are meaningless on LP but meaningful on the more tolerant PR.

In addition, PR isn't self-refuting. LP says "Any untestable statement is without meaning." But LP itself is untestable, and so is meaningless on its own terms. But PR says "Any statement whose truth or falsity could make no conceivable practical difference is without meaning." PR makes a practical difference to our choices about beliefs; presumably we shouldn't believe statements that fail the PR test. And so PR can be meaningful on its own terms.

So we can explain words by definitions, examples, verification conditions, and practical differences. Another way to convey meaning is by contextual use: we 0043 use a word in such a way that its meaning can be gathered from surrounding "clues." Suppose a person getting in a car says "I'm getting in my C"; we can surmise that C means "car." We all learned language mostly by picking up meaning from contextual use.

Some thinkers want us to pick up their technical terms in this same way. We are given no definitions of key terms, no examples to clarify their use, and no explanations in terms of verification conditions or practical differences. We are just told to dive in and catch the lingo by getting used to it. We should be suspicious of this. We may catch the lingo, but it may turn out to be empty and without meaning. That's why the positivists and pragmatists emphasized finding the "cash value" of ideas in terms of verification conditions or practical differences. We must be on guard against empty jargon.

3.4a Exercise

Would each claim be meaningful or meaningless on LP? (Take LP to require that a statement be publicly testable.) Would each be meaningful or meaningless on PR?

Unless we have strong reasons to the contrary, we ought to believe what sense experience seems to reveal.
This is meaningless on LP, since claims about what one ought to do aren't publicly testable. It's meaningful on PR, since its truth could make a difference about what choices we ought to make about beliefs.

1. It's cold outside.

2. That clock is fast.

3. There are five-foot-long blue ants in my bedroom.

4. Nothing is real.

5. Form is metaphysically prior to matter.

6. At noon all lengths, distances, and velocities in the universe will double.

7. I'm wearing an invisible hat that can't be felt or perceived in any way.

8. Regina has a pain in her little toe but shows no signs of this and will deny it if you ask her.

9. Other humans have no thoughts or feelings but only act as if they do.

10. Manuel will continue to have conscious experiences after his physical death.

11. Angels exist (that is, there are thinking creatures who have never had spatial dimensions or weights).

12. God exists (that is, there's a very intelligent, powerful, and good personal creator of the universe).

13. One ought to be logically consistent.

14. Any statement whose truth or falsity could make no conceivable practical difference is meaningless. (PR)

15. Any statement that isn't observationally testable is meaningless. (LP) 0044

3.5 Making distinctions

Philosophers faced with difficult questions often make distinctions:

> "If your question means … [giving a clear phrasing], then my answer is …. But if you're really asking …, then my answer is …."

The ability to formulate various possible meanings of a question is a valuable skill. Many of the questions that confront us are vague or confused; we often have to clarify a question before we can answer it intelligently. Getting clear on a question can be half the battle.

Consider this question (in which I underlined the tricky word "indubitable"):

> "Are some beliefs **indubitable**?"

What does "indubitable" here mean? Does it mean not actually doubted? Or psychologically impossible to doubt? Or irrational to doubt? And what is it to doubt? Is it to refrain from believing? Or is it to have some suspicion about the belief (although we might still believe it)? And indubitable by whom? By everyone (even crazy people)? By all rational persons? By at least some individuals? By me? Our little question hides a sea of ambiguities. Here are three of the many things that our little question could be asking:

- Are there some beliefs that no one has ever refused to believe? (To answer this, we'd need to know whether people in insane asylums sometimes refuse to believe that they exist or that "2 = 2.")
- Are there some beliefs that no rational person has suspicions about? (To answer this, we'd first have to decide what we mean by "rational.")
- Are there some beliefs that some specific individuals are psychologically unable to have any doubts about? (Perhaps many are unable to have any doubts about what their name is or where they live.)

It's risky to answer questions that we don't understand.

Unnoticed ambiguities can block communication. Often people are unclear about what they're asking, or take another's question in an unintended sense. This is more likely if the discussion goes abstractly, without examples.

3.5a Exercise

Each of the following questions is obscure or ambiguous as it stands. Distinguish at least three interesting senses of each question. Formulate each sense simply, clearly, and briefly – and without using the underlined words. 0045

Can one **prove** that there are external objects?

- Can we deduce, from premises expressing immediate experience (like "I seem to see a blue shape"), that there are external objects?
- Can anyone give an argument that will convince (all or most) skeptics that there are external objects?
- Can anyone give a good deductive or inductive argument, from premises expressing their immediate experience in addition to true principles of evidence, to conclude that it's reasonable to believe that there are external objects? (These "principles of evidence" may include things like "Unless we have strong reasons to the contrary, it's reasonable to believe what sense experience seems to reveal.")

1. Is ethics a **science**?
2. Is this monkey a **rational** animal?
3. Is this belief part of **common sense**?
4. Are material objects **objective**?
5. Are values relative (or **absolute**)?
6. Are scientific generalizations ever **certain**?
7. Was the action of that monkey a **free** act?
8. Is **truth** **changeless**?
9. How are moral beliefs **explainable**?
10. Is that judgment based on **reason**?
11. Is a fetus a **human being** (or **human person**)?
12. Are values **objective**?
13. What is the **nature** of **man**?
14. Can I ever **know** what someone else feels?
15. Do you have a **soul**?
16. Is the world **illogical**?

3.6 Analytic and synthetic

Immanuel Kant long ago introduced two related distinctions that have become influential. He divided statements, on the basis of their meaning, into *analytic* and *synthetic* statements. He divided knowledge, on the basis of how it's known, into *a priori* and *a posteriori* knowledge. We'll consider these distinctions in this section and the next.[1]

Kant gave two definitions of "analytic statement":

1. An *analytic statement* is one whose subject contains its predicate.
2. An *analytic statement* is one that's self-contradictory to deny. 0046

Consider these examples (and take "bachelor" to mean "unmarried man"):

(a) "All bachelors are unmarried."
(b) "If it's raining, then it's raining."

Both examples are analytic by definition 2, since both are self-contradictory to deny. But only (a) is analytic by definition 1. In (a), the subject "bachelor" ("unmarried man") contains the predicate "unmarried"; but in (b), the subject "it" doesn't contain the predicate.

We'll adopt definition 2; so we define an **analytic statement** as one that's self-contradictory to deny. *Logically necessary truth* is another term for the same idea; such truths are based on logic, the meaning of concepts, or necessary connections between properties. Here are some further analytic statements:

"2 = 2"
"1 > 0"
"All frogs are frogs."
"If everything is green, then this is green."
"If there's rain, then there's precipitation."
"If this is green, then this is colored."

By contrast, a **synthetic statement** is one that's neither analytic nor self-contradictory; *contingent* is another term for the same idea. Statements divide into analytic, synthetic, and self-contradictory; here's an example of

[1] I'll sketch a standard approach to these Kantian distinctions. Willard Quine, in his *Philosophy of Logic*, 2nd ed. (Cambridge, Mass.: Harvard University Press, 1986), criticizes this approach.

each:[1]

Analytic: "All bachelors are unmarried."
Synthetic: "Daniel is a bachelor."
Self-contradictory: "Daniel is a married bachelor."

While there are three kinds of statement, there are only two kinds of truth: analytic and synthetic. Self-contradictory statements are necessarily false.

3.6a Exercise

Say whether each of these is analytic or synthetic. Take terms in their most natural senses. Some examples are controversial.

All triangles are triangles.
This is analytic. It would be self-contradictory to deny it and say "Some triangles aren't triangles."

1. All triangles have three angles.
2. 2 + 2 = 4.
3. Combining two drops of mercury with two other drops results in one big drop.
4. There are ants that have established a system of slavery.
5. Either some ants are parasitic or else none are.
6. No three-year-old is an adult. 0047
7. No three-year-old understands symbolic logic.
8. Water boils at 90ºC on that 10,000-foot mountain.
9. Water boils at 100ºC at sea level.
10. No uncle who has never married is an only child.
11. All swans are white.
12. Every material body is spatially located and has spatial dimensions.
13. Every material body has weight.
14. The sum of the angles of a Euclidian triangle equals 180º.
15. If all Parisians are French and all French are European, then all Parisians are European.

[1] Modal logic (Chapters 10 and 11) symbolizes "A is analytic (necessary)" as "☐A," "A is synthetic (contingent)" as "(◇A • ◇~A)," and "A is self-contradictory" as "~◇A."

16. Every event has a cause.

17. Every effect has a cause.

18. We ought to treat a person not simply as a means but always as an end in itself.

19. One ought to be logically consistent.

20. God exists.

21. Given that we've observed that the sun rose every day in the past, it's reasonable for us to believe that the sun will rise tomorrow.

22. Unless we have strong reasons to the contrary, we ought to believe what sense experience seems to reveal.

23. Everything red is colored.

24. Nothing red is blue (at the same time and in the same part and respect).

25. Every synthetic statement that's known to be true is known on the basis of sense experience. (There's no synthetic *a priori* knowledge.)

3.7 *A priori* and *a posteriori*

Philosophers traditionally distinguish two kinds of knowledge. ***A posteriori*** (**empirical**) **knowledge** is based on sense experience. ***A priori*** (**rational**) **knowledge** is based on reason, not sense experience. Here's an example of each:

A posteriori: "Some bachelors are happy."
A priori: "All bachelors are unmarried."

While we know both to be true, *how* we know them differs. We know the first statement from our experience of bachelors; we've met many bachelors and recall that some have been happy. If we had to justify the truth of this statement to others, we'd appeal to experiential data about bachelors. In contrast, we know the second statement by grasping what it means and seeing that it must be true. If we had to justify the truth of this statement, we wouldn't have to gather experiential data about bachelors.

Most knowledge is *a posteriori* – based on sense experience. "Sense experience" here covers the five "outer senses" (sight, hearing, smell, taste, and touch). It also covers "inner sense" (the awareness of our own thoughts and feelings) and any other experiential access to the truth that we might have (perhaps even 0048 mystical experience or extrasensory perception).

Logical and mathematical knowledge is generally *a priori*. To test the validity of an argument, we don't go out and do experiments. Instead, we just think and reason; sometimes we write things out to help our thinking. The validity tests in this book are rational (*a priori*) methods. "Reason" in a narrow sense (in which it contrasts with "experience") deals with what we can know *a priori*.

A priori knowledge requires some experience. We can't know that all bachelors are unmarried unless we've learned the concepts involved; this requires experience of language and of (married and unmarried) humans. And knowing that all bachelors are unmarried requires the experience of thinking. So *a priori* knowledge depends somewhat on experience (and thus isn't just something that we're born with). But it still makes sense to call such knowledge *a priori*. Suppose we've gained the concepts using experience. Then to justify the claim that all bachelors are unmarried, we don't have to appeal to any further experience, other than thinking. In particular, we don't have to investigate bachelors to see whether they're all unmarried.[1]

Here are some further examples of statements known *a priori*:

"2 = 2"
"1 > 0"
"All frogs are frogs."
"If everything is green, then this is green."
"If there's rain, then there's precipitation."
"If this is green, then this is colored."

We also gave these as examples of analytic statements.

So far, we've used only analytic statements as examples of *a priori* knowledge and only synthetic statements as examples of *a posteriori* knowledge. Some philosophers think there's only one distinction, but drawn in two ways:

a priori knowledge = analytic knowledge
a posteriori knowledge = synthetic knowledge 0049

Is this view true? If it's true at all, it's not true just because of how we defined the terms. By our definitions, the basis for the analytic / synthetic distinction differs from the basis for the *a priori* / *a posteriori* distinction. A statement is analytic or synthetic depending on whether its denial is self-contradictory; but knowledge is *a posteriori* or *a priori* depending on whether it rests on sense experience. Our definitions leave it open whether the two distinctions coincide.

[1] David Hume, who thought that all concepts come from experience, defended *a priori* knowledge. By comparing two empirical concepts, we can sometimes recognize that the empirical conditions that would verify one ("bachelor") would also verify the other ("unmarried"); so by reflecting on our concepts, we can see that all bachelors must be unmarried.

These two combinations are very common:

analytic *a priori* knowledge
synthetic *a posteriori* knowledge

Most of our knowledge in math and logic is analytic *a priori*. Most of our everyday and scientific knowledge about the world is synthetic *a posteriori*. These next two combinations are more controversial:

analytic *a posteriori* knowledge
synthetic *a priori* knowledge

Can we know any analytic statements *a posteriori*? It seems that we can. "π is a little over 3" is presumably an analytic truth that can be known either by *a priori* calculations (the more precise way to compute π) – or by measuring circles empirically (as the ancient Egyptians did). And "It's raining or not raining" is an analytic truth that can be known either *a priori* (and justified by truth tables, see §6.6) – or by deducing it from the empirical statement "It's raining." But perhaps any analytic statement that's known *a posteriori* also could be known *a priori*.

The biggest issue is this: "Do we have any synthetic *a priori* knowledge?" This asks whether there's any statement A such that:

- A is synthetic (not self-contradictory either to affirm or to deny),
- we know A to be true, and
- our knowledge of A is based on reason (and not sense experience)?

In one sense of the term, an *empiricist* is one who rejects such knowledge – and who thus limits what we can know by pure reason to analytic statements. By contrast, a *rationalist* is one who accepts such knowledge – and who thus gives a greater scope to what we can know by pure reason.[1]

Empiricists deny the possibility of synthetic *a priori* knowledge for two main reasons. First, it's difficult to understand how there could be such knowledge. Analytic *a priori* knowledge is fairly easy to grasp. Suppose a statement is true simply because of the meaning and logical relations of the concepts involved; then we can know it in an *a priori* fashion by reflecting on these concepts and logical relations. But suppose a statement could logically be either true or false. How could we then possibly know by pure thinking which it is?

Second, those who accept synthetic *a priori* truths differ on what these

[1] More broadly, *empiricists* are those who emphasize *a posteriori* knowledge, while *rationalists* are those who emphasize *a priori* knowledge.

truths are. They just follow their prejudices and call them “deliverances of reason.”

Rationalists accept synthetic *a priori* knowledge for two main reasons. First, the opposite view (at least if it’s claimed to be *known*) seems self-refuting. Consider empiricists who claim to know *“There’s no synthetic a priori knowledge.”* Now this claim is synthetic (it’s not true by how we defined the terms “synthetic” and “*a priori*,” and it’s not self-contradictory to deny). And it would have to be known *a priori* (since we can’t justify it by sense experience). So the empiricist’s claim would have to be synthetic *a priori* knowledge, which it rejects. 0050

Second, we seem to have synthetic *a priori* knowledge of ideas like this:

> If you believe you see an object to be red and have no special reason to doubt your perception [e.g., the lighting is strange or you’re on mind-altering drugs], then it’s reasonable for you to believe that you see an actual red object.

This claim is synthetic; it’s not true because of how we’ve defined terms, and skeptics can deny it without self-contradiction. It’s presumably known to be true; if we didn’t know such truths, then we couldn’t justify any empirical beliefs. And it’s known *a priori*; empirical knowledge depends on it instead of it depending on empirical knowledge. So we have synthetic *a priori* knowledge of this claim. So there’s synthetic *a priori* knowledge.

The dispute over synthetic *a priori* knowledge influences how we do philosophy. Can basic ethical principles be known *a priori*? Empiricists say no; so then we know basic ethical principles either empirically or not at all. But rationalists can (and often do) think that we know basic ethical truths *a priori*, from reason alone (through either intuition or some rational consistency test).

3.7a Exercise

Suppose we knew each of these to be true. Would our knowledge likely be *a priori* or *a posteriori*? Take terms in their most natural senses. Some examples are controversial.

All triangles are triangles.
This would be known *a priori*.

Use the examples from §3.6a.

4 Fallacies and Argumentation

This chapter deals with arguing well, recognizing fallacies, avoiding inconsistencies, developing your own arguments, and analyzing arguments that you read.

4.1 Good arguments

A **good argument**, to be logically correct and fulfill the purposes for which we use arguments, should:

1. be deductively valid (or inductively strong) and have premises all true;
2. have its validity (or inductive strength) and truth-of-premises be as evident as practically possible to the parties involved;
3. be clearly stated (using understandable language and making clear what the premises and conclusion are);
4. avoid circularity, ambiguity, and emotional language; and
5. be relevant to the issue at hand.

If you fulfill these, then you're arguing well.

First, a good argument should be deductively valid (or inductively strong – see Chapter 5) and have premises all true. We often criticize an argument by trying to show that the conclusion doesn't follow from (or isn't supported by) the premises, or that one or more of the premises are false.

Second, a good argument should have its validity (or inductive strength) and truth-of-premises be as evident as practically possible to the parties involved. Arguments are less effective if they presume premises that others see as false or controversial. Ideally, we'd like to use only premises that everyone will accept as immediately obvious; but in practice, this is too high an ideal. We often appeal to premises that will only be accepted by those of similar political, religious, or philosophical views. And sometimes we appeal to hunches, like "I can get to the gun before the thief does"; while not ideal, this may be the best we can do at a given moment.

Third, a good argument should be clearly stated; it should use understandable language and make clear what the premises and conclusion are. Obscure or overly complex language makes reasoning harder to grasp.

When we develop an argument, a good strategy is to put it on paper in a 0052 preliminary way and then reread it several times, trying to make improvements. Try to express the ideas more simply and clearly, and think how others may object or misunderstand. Often ideas first emerge in a confused form; clarity comes later, after much hard work. While mushy thinking is often unavoidable in the early development of an idea, it's not acceptable in the final product.

People often argue without making clear what their premises and conclusions are; sometimes we get stream-of-consciousness ramblings sprinkled with an occasional "therefore." While this is unacceptable, a good argument needn't spell everything out; it's often fine to omit premises that are obvious to the parties involved. If I'm hiking on the Appalachian Trail, I might say to my hiking partner: "We can't still be on the right trail, since we don't see white blazes on the trees." This is fine if my partner knows that we'd see white blazes if we were on the right trail; then the full argument would be pedantic:

We don't see white blazes on the trees.
If we were still on the right trail, then we'd see white blazes on the trees.
∴ We aren't still on the right trail.

In philosophy, it's often wise to spell out *all* our premises, since unstated ideas are often crucial but unexamined. Suppose someone argues: "We can't be free, since all our actions are determined." This assumes the italicized premise:

All human actions are determined.
No determined action is free.
∴ No human actions are free.

We should be aware that we're assuming this controversial premise.

So a *good argument* should be valid (or inductively strong) and have premises all true, this validity/strength and truth should be as evident as practically possible, and it should be clearly stated. Our final conditions say that the argument should (4) avoid circularity, ambiguity, and emotional language; and (5) be relevant to the issue at hand. Five common fallacies tie into these final conditions.

Our first fallacy is *circularity*:

An argument is **circular** if it presumes the truth of what is to be proved.

A series of arguments is **circular** if it uses a premise to prove a conclusion – and then uses that conclusion to prove the premise.

"The soul is immortal because it can't die" is circular; since the premise just repeats the conclusion in different words, the argument takes for granted what it's supposed to prove. A circular *series* of arguments might say: "A is true because B is true, and B is true because A is true." A circular argument is also said to be *question begging*; this differs from the new (and confusing) usage in 0053 which "begging a question" means "raising a question."

Here's a second fallacy, and a crude argument that exemplifies it:

An argument is **ambiguous** if it changes the meaning of a term or phrase within the argument.

Love is an emotion.
God is love.
∴ God is an emotion.

Premise 1 requires that we take "love" to mean "the feeling of love" – which makes premise 2 false or doubtful. Premise 2 requires that we take "love" to mean "a supremely loving person" or "the source of love" – which makes premise 1 false or doubtful. We can have both premises clearly true only by shifting the meaning of "love." Ambiguity is also called *equivocation*.

Unclear sentence structures can bring ambiguities. For example, "pretty little girls' camp" can mean "camp for little girls who are pretty," "pretty camp for little girls," or "pretty camp that is little and for girls."

It's important to avoid emotionally slanted terms when we reason:

To **appeal to emotion** is to stir up feelings instead of arguing in a logical manner.

Students, when asked to argue against a theory, often just describe the theory in derogatory language; so a student might dismiss Descartes by calling his views "superficial" or "overly dualistic." But such verbal abuse doesn't give any reason for thinking a view wrong. Often the best way to argue against a theory is to find some false implication and then reason as follows:

If the theory is true, then this other thing also would be true.
This other thing isn't true.
∴ The theory isn't true.

Recall that an argument consists of premises and a conclusion.

Our last condition says that a good argument must be relevant to the issue at hand. A clearly stated argument might prove something and yet still be defective, since it may be beside the point in the current context:

An argument is **beside the point** if it argues for a conclusion irrelevant to the issue at hand.

Hitler, when facing a group opposed to the forceful imposition of dictatorships, shifted their attention by attacking pacifism; his arguments, even if sound, were beside the point. Such arguments are also called *red herrings*, after a practice used in training hunting dogs: a red herring fish would be dragged across the trail to 0054 distract the dog from tracking an animal. In arguing, we must keep the point at issue clearly in mind and not be misled by a smelly fish.

Students sometimes use this "beside the point" label too broadly, to apply to almost any fallacy. This fallacy isn't about the *premises being irrelevant* to the conclusion; instead, it's about the *conclusion* (regardless of whether it's proved) *being irrelevant to the issue at hand*. Suppose a politician is asked "Where do you stand about the proposed tax cuts?" but evades answering, instead shifting our attention to the need for a strong military. These statements are *beside the point*, since they don't answer the question.

One common form of this fallacy has its own name:

A **straw man** argument misrepresents an opponent's views.

This is common in politics. Candidate A for mayor suggests cutting a few seldom-used stations on the rapid transit system. Then candidate B's campaign ad expresses shock that A wants to dismantle the whole transit system, which so many citizens depend on; the ad attacks, not what A actually holds, but only a "straw man" – a scarecrow of B's invention. Campaign ads and speeches that distort an opponent's view have recently got so bad that "fact checkers" and "truth squads" have arisen to point out misleading language and downright falsehoods – regardless of which side engages in these.

Again, a *good argument* is valid (or inductively strong) and has premises all true; has this validity/strength and truth be as evident as practically possible; is clearly stated; avoids circularity, ambiguity, and emotional language; and is relevant to the issue. Good arguments normally convince others, but not always. Some people aren't open to rational argument on many issues; some believe that the earth is flat, despite good arguments to the contrary. And bad arguments sometimes convince people; Hitler's *beside the point* fallacy and the candidate's *straw man* fallacy can mislead and convince. Studying logic can help protect us from bad reasoning. The better we can distinguish good from bad reasoning, the less will politicians and others be able to manipulate people.

"Proof" is roughly like "good argument." But we can *prove* something even if our argument is unclear, contains emotional language, or is irrelevant to the issue at hand. And a proof must be *very* strong in its premises and in how it connects the premises to the conclusion; for the latter reason, it seems wrong to call inductive arguments "proofs." So we can define a **proof** as a non-circular, non-ambiguous, deductively valid argument with clearly true premises. A **refutation** of a statement is a proof of the statement's denial.

"Proof" can have other meanings. Chapters 7 to 14 use "proof" in the technical sense of "formal proof," to cover logical derivations that follow specified rules. And Exercise 3.5a explained that "prove" could have various meanings in the question, "Can we prove that there are external objects?" The word "proof" has a cluster of related meanings. 0055

"Prove" and "refute" are often misused. These properly apply only to successful arguments. A *proof* shows that something is true, and a *refutation* shows that something is false. Avoid saying things like "Hume proved this, but Kant refuted him." This is self-contradictory, since it implies that Hume's claim is both true and false – and that Hume showed it was true and Kant showed it was false. It's better to say "Hume argued for this, but Kant criticized his reasoning."

4.2 Informal fallacies

A **fallacy** is a deceptive error of thinking; an **informal fallacy** is a fallacy not covered by some system of deductive or inductive logic. In working out the conditions for a good argument, we introduced five informal fallacies: *circular*, *ambiguous*, *appeal to emotion*, *beside the point*, and *straw man*. We now add thirteen more, in three groups. While this book covers most common fallacies, there are many others that aren't listed here.

Our first group includes six fallacies expressed in a premise–conclusion format. This first appeals to our herd instincts:

Appeal to the crowd
Most people believe A. ∴ A is true.
Most people think Wheaties is very nutritious. ∴ Wheaties is very nutritious.

Despite what people think, Wheaties cereal might have little nutritional value. Discovering its nutritional value requires checking its nutrient content; group opinion proves nothing. While we all recognize the fallacy here, group opinion still may influence us. Humans are only partially rational.

The *opposition fallacy* comes from dividing people into “our group” (which has the truth) and “our opponents” (who are totally wrong):

Opposition
Our opponents believe A. ∴ A is false.
Those blasted liberals say we should raise taxes. ∴ We shouldn’t raise taxes.

The problem here is that our opponents may be right.

The *genetic fallacy* dismisses a belief on the basis of its origin:

Genetic fallacy
We can explain why you believe A. ∴ A is false.
Any psychologist would see that you believe A because of such and such. ∴ A is false.

One who has superficially studied a little psychology may dismiss another’s
0056 views in this way. An appropriate (but nasty) reply is, “And what is the psychological explanation for why you confuse psychological explanations with logical disproofs?” To show a belief to be false, we must argue against the *content* of the belief; it’s not enough to explain how the belief came to be.

This next one has two closely related forms:

Appeal to ignorance
No one has proved A. ∴ A is false. No one has disproved A. ∴ A is true.
No one has proved there's a God. ∴ There's no God. No one has proved there's no God. ∴ There's a God.

Something not proved might still be true, just as something not disproved might still be false. An "appeal to ignorance" must have one of these forms; it's not just any case where someone speaks out of ignorance.

This next one uses a Latin name for "after this therefore because of this":

Post hoc ergo propter hoc
A happened after B. ∴ A was caused by B.
Paul had a beer and then got 104% on his logic test. ∴ He got 104% because he had beer.

The premise was true (there were bonus points). Some students concluded: "So if I have a beer before the test, I'll get 104%" and "If I have a six-pack, I'll get 624%." Proving causal connections requires more than just the sequence of two factors; the factors might just *happen* to have occurred together. It's not even enough that factors *always* occur together; day always follows night, and night always follows day, but neither causes the other. Proving causal connections is difficult (see Mill's methods in §5.7).

This next one is also called *division–composition*:

Part–whole
This is F. ∴ Every part of this is F. Every part of this is F. ∴ This is F.
My essay is good. ∴ Every sentence of my essay is good. Every sentence of my essay is good. ∴ My essay is good.

The first argument is wrong because an essay might be good despite having some poor sentences. The second is wrong because each sentence of the essay might be good without the essay as a whole being good; the fine individual sentences might not make sense together. So something might be true of a whole without being true of the parts, or true of the parts without being true of the whole. A property that characterizes a whole but not any parts is sometimes called an *emergent property*: being alive is an emergent property possessed by a cell but not by any component molecules – and water may be clear and wet without 0057 individual H_2O molecules being clear or wet. More controversially, some say thinking is an emergent property possessed by the brain but not by its cells.

In rare cases, these fallacy forms might be abbreviated forms of good reasoning. Suppose you know that people in your society almost never have false beliefs; then this "appeal to the crowd" could be correct inductive reasoning:

Almost always, what most people in my society believe is true.
Most people in my society believe A.
That's all we know about the matter.
∴ Probably A is true.

Or suppose you know that your opponent Jones is always wrong. Then this could be sound reasoning: "Everything Jones says is false, Jones says A, so A is false." But correct forms of these six fallacy forms are unusual in real life.

Our next group has three types of reasoning with correct and fallacious forms. This first type of reasoning appeals to expert opinion:

Appeal to authority – correct form:

X holds that A is true.
X is an authority on the subject.
The consensus of authorities agrees with X.
∴ There's a presumption that A is true.

Incorrect forms omit premise 2 or 3, or conclude that A *must* be true.

This one has the correct form:

Your doctor tells you A.
She's an authority on the subject.
The other authorities agree with her.
∴ There's a presumption that A is true.

This conclusion means that we ought to believe A unless we have special evidence to the contrary. If the doctor is a great authority and the consensus of authorities is large, then the argument becomes stronger; but it's never totally conclusive. All the authorities in the world might agree on something that they later discover to be wrong; so we shouldn't think that something *must* be so because the authorities say it is. It's also wrong to appeal to a person who isn't an authority in the field (a sports hero endorsing coffee makers, for example). And finally, it's weak to appeal to one authority (regarding the safety of nuclear energy, for example) when the authorities disagree widely. The appeal to authority can go wrong in many ways. Yet many of our trusted beliefs (that Washington was the first US president, for example, or that there's such a country as Japan) rest quite properly on the say so of others.

An "authority" might be a calculator or computer instead of a human. My calculator has proved itself reliable, and it gives the same result as other reliable calculators. So I believe it when it tells me that $679 \bullet 177 = 120{,}183$.

0058

This next one uses a Latin name for "against the person" (which is opposed to *ad rem*, "on the issue"):

Ad hominem – correct form:

X holds that A is true.
In holding this, X violates legitimate rational standards (for example, X is inconsistent, biased, or not correctly informed).
∴ X isn't fully reasonable in holding A.

Incorrect forms use factors irrelevant to rational competence (for example, X is a member of a hated group or beats his wife) or conclude that A is false.

This one has the correct form: Rick holds that people of this race ought to be treated poorly. In holding this, Rick is inconsistent (because he doesn't think that he ought to be treated that way if he were in their exact place) and so violates legitimate rational standards. ∴ Rick isn't fully reasonable in his views.

A "personal attack" argument can be either legitimate or fallacious. In our example, we legitimately conclude that Rick, because he violates rational standards, isn't fully reasonable in his beliefs. It would be fallacious to draw the stronger conclusion that his beliefs must be wrong; to show his beliefs to be wrong, we must argue against the beliefs, not against the person. A more extreme *ad hominem* was exemplified by Nazis who argued that Einstein's theories must be wrong since he was Jewish; being Jewish was irrelevant to Einstein's competence as a scientist.

This next form of reasoning lists and weighs reasons for and against:

Pro–con – correct form: The reasons in favor of act A are …. The reasons against act A are …. The former reasons outweigh the latter. ∴ Act A ought to be done.
Incorrect form: The reasons in favor of act A are …. ∴ Act A ought to be done.
This one has the correct form: The reasons in favor of getting an internal-frame backpack are …. The reasons against getting an internal-frame backpack are …. The former reasons outweigh the latter. ∴ I ought to get an internal-frame backpack.

People sometimes make decisions by folding a piece of paper in half and listing *reasons in favor* on one side and *reasons against* on the other; then they decide intuitively which side has stronger (not necessarily more) reasons. This method forces us to look at both sides of an issue. In the incorrect form, we just look at half the picture; we say that you should do this (because of such and such advantages) or that you shouldn't do it (because of such and such disadvantages). 0059 This fallacy is also called "one-sided" or "stacking the deck."

We can expand our three correct forms into standard inductive and deductive arguments. A correct appeal to authority becomes a strong inductive argument if we add this inductively confirmed premise: "The consensus of authorities on a subject is usually right." Correct *ad hominem* arguments become deductively valid if we add: "Anyone who, in believing A, violates legitimate rational standards is thereby not fully reasonable in believing A." And correct pro–con arguments become deductively valid if we add: "If the reasons in favor of A outweigh the reasons against A, then A ought to be done."

Our final group has four miscellaneous fallacies. Here's the first fallacy (which is also called *false dilemma*):

> **Black-and-white** thinking oversimplifies by assuming that one or another of two extreme cases must be true.

One commits this fallacy in thinking that people must be *logical* or *emotional*, but can't be both. My thesaurus lists these terms as having opposite meanings; but if they really had opposite meanings, then no one could be both at once – which indeed is possible. In fact, all four combinations are common:

logical and emotional
logical and unemotional
illogical and emotional
illogical and unemotional

People who think in a black-and-white manner prefer simple dichotomies, like logical-emotional, capitalist-socialist, or intellectual-jock. Such people have a hard time seeing that the world is more complicated than that.

This next fallacy is also called *hasty generalization*:

> To use a **false stereotype** is to assume that the members of a certain group are more alike than they actually are.

People commit this fallacy in thinking that all Italians exist only on spaghetti, that all New Yorkers are uncaring, or that all who read Karl Marx want to overthrow the government. False stereotypes can be detrimental to the stereotyped. A study compared scores on a math test of two otherwise identical groups of young girls; just the first group was told beforehand that girls are genetically inferior in math – and this group did much worse on the test.

This next fallacy substitutes violence for reasoning:

To **appeal to force** is to use threats or intimidation to get a conclusion accepted.

0060 A parent might say, “Just agree and shut up!” Parents and teachers hold inherently intimidating positions and are often tempted to appeal to force.

This last fallacy is also called *trick question*:

A **complex question** is a question that assumes the truth of something false or doubtful.

The standard example is: “Are you still beating your wife?” A “yes” implies that you still beat your wife, while a “no” implies that you used to beat her. The question combines a statement with a question: “You have a wife and used to beat her; do you still beat her?” The proper response is: “Your question presumes something that’s false, namely that I have a wife and used to beat her.” Sometimes it’s misleading to give a “yes” or “no” answer.

4.2a Exercise: LogiCola R

Identify the fallacies in the following examples. Not all are clear-cut; some examples are controversial and some commit more than one fallacy. All the examples here are fallacious. Use these labels to identify the fallacies:

aa = appeal to authority
ac = appeal to the crowd
ae = appeal to emotion
af = appeal to force
ah = *ad hominem*
ai = appeal to ignorance
am = ambiguous
bp = beside the point
bw = black and white
ci = circular
cq = complex question
fs = false stereotype
ge = genetic
op = opposition
pc = pro–con
ph = *post hoc*
pw = part–whole
sm = straw man

This sports hero advertises a popcorn popper on TV. He says it's the best popcorn popper, so this must be true.
This is aa (appeal to authority). There's no reason to think the sports hero is an authority on popcorn poppers.

1. Are you still wasting time with all that book-learning at the university?

2. The Bible tells the truth because it's God's word. We know the Bible is God's word because the Bible says so and it tells the truth.

3. You should vote for this candidate because she's intelligent and has much experience in politics.

4. The Equal Rights Amendment was foolish because its feminist sponsors were nothing but bra-less bubbleheads.

5. No one accepts this theory anymore, so it must be wrong.

6. Either you favor a massive arms buildup, or you aren't a patriotic American.

7. The president's veto was the right move. In these troubled times we need decisive leadership, even in the face of opposition. We should all thank the president for his courageous move.

8. Each member of this team is unbeatable, so this team must be unbeatable. 0061

9. My doctor told me to lose weight and give up smoking. But she's an overweight smoker herself, so I can safely ignore her advice.

10. Belief in God is explained in terms of one's need for a father figure; so it's false.

11. There are scientific laws. Where there are laws there must be a lawgiver. Hence someone must have set up the scientific laws to govern our universe, and this someone could only be God.

12. The lawyer for the defense claims that there's doubt that Smith committed the crime. But, I ask, are you going to let this horrible crime go unpunished because of this? Look at the crime; see how horrible it was! So you see clearly that the crime was horrible and that Smith should be convicted.

13. Free speech is for the common good, since unrestrained expression of opinion is in people's interest.

14. This is a shocking and stupid proposal. Its author must be either a dishonest bum or a complete idiot.

15. Aristotle said that heavy objects fall faster than light ones, so it must be true.

16. Each of these dozen cookies (or drinks) by itself isn't harmful; one little one won't hurt! Hence having these dozen cookies (or drinks) isn't harmful.

17. Before Barack Obama became the Democratic candidate for US president, he ran in a series of primary elections. He noted that he played basketball before the Iowa primary, and then won the vote, while he neglected to play before the New Hampshire primary, and then lost. He concluded (in jest) "At that point I was certain that we had to play on every primary."

18. Only men are rational animals. No woman is a man. Therefore no woman is a rational animal.

19. I'm right, because you flunk if you disagree with me!

20. The discriminating backpacker prefers South Glacier tents.

21. Those who opposed the war were obviously wrong; they were just a bunch of cowardly homosexual Communists.

22. We should legalize gambling in our state, because it would bring in new tax revenue, encourage tourists to come and spend money here, and cost nothing (just the passing of a new law).

23. Do you want to be a good little boy and go to bed?

24. This man is probably a Communist. After all, nothing in the files disproves his Communist connections.

25. People who read *Fortune* magazine make a lot of money. So if I subscribe to *Fortune*, then I too will make a lot of money.

26. Feminists deny all difference between male and female. But this is absurd, as anyone with eyes can see.

27. Each part of life (eyes, feet, and so on) has a purpose. Hence life itself must have a purpose.

28. So you're a business major? You must be one of those people who care only about the almighty dollar and aren't concerned about ideas.

29. My opponent hasn't proved that I obtained these campaign funds illegally. So we must conclude that I'm innocent.

30. Those dirty Communists said that we Americans should withdraw from the Panama Canal, so obviously we should have stayed there. 0062

31. Karl Marx was a personal failure who couldn't even support his family, so his political theory must be wrong.

32. Religion originated from myth (which consists of superstitious errors). So religion must be false.

33. Suzy brushed with Ultra Brilliant and then attracted boys like a magnet! Wow – I'm going to get some Ultra Brilliant. Then I'll attract boys too!

34. Did you kill the butler because you hated him or because you were greedy?

35. My parents will be mad at me if I get a D, and I'll feel so stupid. Please? You know how I loved your course. I surely deserve at least a C.

36. Miracles are impossible because they simply can't happen.

37. I figure that a person must be a Communist if he doesn't think the American free-enterprise system is flawless and the greatest system in the world.

38. Everyone thinks this beer is simply the best. So it must be the best.

39. We ought to oppose this, since it's un-American.

40. Practically every heroin addict first tried marijuana. Therefore, marijuana causes heroin addiction.

41. Most college students are mainly concerned with sports, liquor, and sex. So this is normal. But Duane is mainly concerned with poetry. So he must be abnormal and thus unhealthy.

42. Each of the things in my backpack is light, so my loaded backpack must be light.

43. You're wrong in disagreeing with me, because what I said is true.

44. Everyone thinks the Democrat is the better candidate, so it must be true.

45. We should reject Mendel's genetic theories, since he was a monk and thus couldn't have known anything about science.

46. Every time I backpack it seems to rain. I'm going backpacking next week. So this will cause it to rain.

47. It hasn't been proved that cigarettes are dangerous, so it's only reasonable to conclude that they aren't dangerous.

48. In a commercial filled with superb scenery, sexy girls, and soft music: "Buy a Ford Mustang – it's a super car!"

49. Atheism is absurd. Atheists deny God because they can't see him. But who has seen electrons either?

50. President George W. Bush was in office for several years, and then the financial crisis occurred in 2008. Therefore the crisis occurred because Bush was in office.

51. Do you support freedom and the unrestricted right to buy weapons?

52. We don't know how the first forms of life could have emerged by natural causes from the primeval chemical soup that covered the earth. So we must assume that they didn't emerge by natural causes; so they had to have had a divine origin.

53. Since no atom in this rock is heavy or green, this rock cannot be heavy or green.

54. That car can't be any good, since it was made in Detroit.

55. All doctors are men with medical degrees. But no woman is a man with a medical degree. Therefore, no woman is a doctor.

56. If you don't keep quiet about our bank's dishonest practices, you're apt to lose your job.

57. A black cat crossed my path, and then later I flunked my logic test. So this proves that black cats are unlucky. 0063

58. Either you respect and agree with your teacher, or you're insolent and don't deserve a good grade.

59. In spite of warnings from lifeguards, my girlfriend went swimming without a worry. She said that she didn't have to worry about man-eating sharks.

60. Will you contribute to our collection for famine relief, or are you insensitive to the suffering of other people?

4.2b Another Fallacy Exercise: LogiCola R

1. When are we going to guarantee all the people of this country the health care that they deserve?

2. When are we going to understand that the government cannot afford to pay for universal health care?

3. The professor's letter of recommendation said, "I cannot praise this student's study habits too highly."

4. No one has proven that humans are causing global warming; so we should assume that the heating of the earth has purely natural causes.

5. Christians are peaceful, Muslims are terrorists.

6. I never had problems with headaches before I studied logic. So studying logic must cause my headaches.

7. This candidate's ideas are really scary; don't they make you afraid? I fear what would happen to our country if this candidate were elected.

8. Charles Darwin, who came up with the theory of evolution, presumably thought that his grandfather was a monkey.

9. You ask me why I deposited the company funds in my personal banking account. But why are you so doubtful about my integrity? Don't you believe that we all need to be more trusting?

10. American military experts testified in the first decade of the 21st century that Iraq was developing weapons of mass destruction; so this must be true.

11. If all persons in a group work to maximize their individual self-interest, then the group is working effectively to maximize its own self-interest.

12. The liberal elite media did it again! Those idiots are out to attack those of us who have solid, pro-American values.

13. My mother demands that I clean up after I make waffles. She is an incredible neatness freak! She wants me to devote my whole life to keeping her kitchen spotless!

14. Liberation theology got some of its concepts (like oppressive social structures) from atheistic Marxists, and so these concepts should be rejected.

15. This backpacking tent is very lightweight, and so this is the one you should get.

16. Everyone knows there ain't no gold in the Grand Canyon.

17. The Democrats want to raise tax rates on the rich and lower them on the middle class. This is part of their plan to move the country into socialism.

18. No one has given conclusive evidence showing that aliens from outside our planet didn't land near Roswell in 1947. So we should believe the witnesses who say that they encountered such aliens.

19. You should vote for me because I will lower your taxes.

20. Humans are "hardwired" so that, at least for the most part, they believe in God. So belief in God is rational. 0064

21. Humans are "hardwired" so that, at least for the most part, they believe in God. So belief in God is irrational.

22. The second exam question asked me to describe Aristotle's approach to ethics. But since I didn't know anything about this, I instead described Plato's approach.

23. Those horrible city folk vote Democratic; so we country folk should vote Republican.

24. If you don't want to suffer an unfortunate accident, you'd better find my client innocent.

25. We should take either all of the Bible literally or else none of literally.

26. Men are logical, women are emotional.

27. Since there's no good evidence that there's intelligent life in other parts of the universe, it's only reasonable to conclude that there's no such life.

28. Since Martin Heidegger developed many of his ideas when he was a Nazi supporter in Germany, we should disregard his ideas.

29. Gensler, who authored the Routledge *Introduction to Logic*, wears sandals with socks and claims that this is very fashionable; so this must be so.

30. We shouldn't listen when this Republican argues for tax relief for the rich; after all, her family was very rich.

31. If you don't buy some Girl Scout Cookies, I'll tell everyone how cheap you are.

32. My favorite Russian tennis star claims that Canon cameras are the best; so I plan to get one.

33. Where did you hide the dead body of your murder victim?

34. I read on the Internet that global warming is a hoax; so this must be true.

35. Cheating on exams can't be wrong; I mean, everyone does it.

36. The Republicans say that they are against "big government." But they really want to eliminate all social services for those in need, so the rich can become even richer.

37. Last night I shot a burglar in my pajamas. I don't know how he got into my pajamas.

38. Are you going to admit that you're wrong?

39. Look at all the bad things that happened to our country while my opponent was in office! If you don't want to elect an official who'll bring about such bad things, then you should vote against my opponent.

40. Everything in the universe has a cause; so the universe also has a cause.

41. If you need another reference for my honesty, I can get Mariana Smith to vouch for me. Oh, you've never heard of Mariana Smith? Well, I can vouch for her.

42. I installed LogiCola on my computer, and then two weeks later my computer failed. LogiCola must be to blame!

43. So, you ask, which of my campaign promises will have to wait if we don't have enough funds to fulfill them all? Instead of responding, I'd like to address what's really troubling the people of this country, namely why the current administration is so dishonest.

44. Either you favor the Republicans or you aren't patriotic.

45. I had foolish and immature ideas like yours when I was your age.

46. Ancient Romans to Christians: "If you refuse to renounce your faith and worship the gods of Rome, we'll feed you to the lions."

47. All logicians are emotionless calculators. 0065

48. When Gensler baked his first batch of cookies, he used very good ingredients. Therefore the cookies that he baked were very good.

49. We shouldn't listen when this Democrat argues for tax relief for the poor; after all, her family was very poor.

50. God must have created the world, since surely *someone* must have created it.

51. Most Americans supported President George W. Bush's invasion of Iraq, so this invasion must have been a good thing.

52. You should take Gensler's logic course because he has a great sense of humor.

53. If you weren't so stupid, you'd agree with me.

54. To a junior Member of Congress: "If you don't vote for this Bill, you'll never be appointed to any important committees."

55. Why does my opponent want to lead our country into socialism?

56. Since each cell in the human organism is incapable of thought, thus the human organism itself is incapable of thought.

57. The Volkswagen was first developed by the Nazis, and so it must be an evil car.

58. Those crude country folk support this idea; so we city folk should be against it.

59. Dr Jones, you can't prove that I didn't come up independently with the same essay that occurs with word-by-word similarity on the Internet. So you must assume that I'm innocent of plagiarism.

60. Gensler's logic book is the best. My proof is that it says so inside, on the last problem of §4.2b.

4.3 Inconsistency

Inconsistency is the most important fallacy – the most important deceptive error of thinking. Students writing on philosophical issues for the first time often express inconsistent views, as in this example:

> Since morality is relative to culture, no duties bind universally. What's right in one culture is wrong in another. Universal duties are a myth. Relativism should make us tolerant toward others; we can't say that we're right and they're wrong. So everyone ought to respect the values of others.[1]

Here the first statement is incompatible with the last:

> "No duties bind universally."
> "Everyone ought to respect the values of others."

If *everyone* ought to respect the values of others, then some duties bind universally. And if *no* duties bind universally, then neither does the duty to respect the values of others. This inconsistency isn't trivial; it cuts deeply. The unexamined views that we use to guide our lives are often radically incoherent; putting these views into words often brings out their incoherence. The ancient Greek philosopher Socrates was adept at showing people how difficult it was to 0066 have consistent beliefs on the deeper questions of life.

Inconsistency is common in other areas too. Someone running for political office might talk to environmentalists one day and industrialists the next. Each group might be told exactly what it wants to hear. The first group is told "I'll support stronger clean-air standards"; the second is told "I'll try to lower clean-air standards." We can be sure that the politician, if elected, will violate some of the promises; one can't fulfill incompatible promises.

We often aren't aware of our inconsistency. For example, one might believe all three of these:

[1] See my *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018), Chapter 2.

1. God is good.
2. Predestination is true. (God immediately causes everything that happens.)
3. God damns sinners to eternal punishment.

These three beliefs aren't inconsistent in themselves. But the person might have other beliefs that add to these three to make an inconsistent set:

4. If predestination is true, then God causes us to sin.
5. If God causes us to sin and yet damns sinners to eternal punishment, then God isn't good.

This set of five beliefs is inconsistent. Beliefs 2 and 4 entail "God causes us to sin." This, with 3 and 5, entails "God isn't good" – which contradicts 1. So the five beliefs can't all be true together. Someone who believes all five might not be aware of the inconsistency; the beliefs might not have come together in the person's consciousness at the same time.

Inconsistency is a sign that our belief system is flawed and that we need to change something. Logic can tell us that our belief system is inconsistent. But it can't tell us how to rearrange beliefs to regain consistency; that's up to us.

Controversies often arise when a set of individually plausible statements can't consistently be combined. Consider this group of statements:

F = Some human actions are free.
D = All human actions are determined.
I = No determined actions are free.

Even though each claim by itself is plausible, the set is inconsistent. If we take any two of the statements as premises, we can infer the denial of the third. *Hard determinists* take D (determinism) and I (that determinism is incompatible with free will) as premises. They conclude not-F (that we have no free will):

All human actions are determined.
No determined actions are free.
∴ No human actions are free.

D
I
∴ Not-F 0067

Indeterminists take F (free will) and I (that determinism is incompatible with free will) as premises. They conclude not-D (the falsity of determinism):

Some human actions are free.
No determined actions are free.
∴ Some human actions aren't determined.

F
I
∴ Not-D

Soft determinists take F (free will) and D (determinism) as premises. They conclude not-I (that determinism isn't incompatible with free will):

Some human actions are free.
All human actions are determined.
∴ Some determined actions are free.

F
D
∴ Not-I

Each of the three arguments has plausible premises. All three arguments are valid, but at most only one of them can have true premises.

The three arguments relate to each other in an interesting way. Each argument is a "turnaround" of the other two. An argument is a *turnaround* of another argument if each results from the other by switching the denial of a premise with the denial of the conclusion. Here is an example:

Hard determinism

D
I
∴ **Not-F**

Indeterminism (switches the denial of a premise with the denial of the conclusion)

F
I
∴ **Not-D**

As you'll see from the exercises, several classical philosophical disputes involve turnaround arguments. In each dispute, we have a set of individually plausible statements that can't consistently be combined.

A single statement may be inconsistent with itself. A **self-refuting statement** is a statement that makes such a sweeping claim that it ends up denying itself. Suppose I tell you: "Everything that I tell you is false." Could this be true? Not if I tell it to you; then it has to be false. The statement refutes itself.

Or suppose I say: "I know that there's no human knowledge." This couldn't be true. If it were true, then there would be some human knowledge – thus refuting the claim. A self-refuting claim often starts as a seemingly big, bold insight. The bubble bursts when we see that it destroys itself.

Consistency relates ethical beliefs to actions in a special way. Suppose I believe that this man is bleeding. That belief doesn't commit me, under pain of inconsistency, to any specific act; how I live can't be inconsistent with this belief (taken by itself). But suppose I believe that I *ought* to call the doctor. This ethical belief does commit me, under pain of inconsistency, to action. If I don't act to call the doctor, then the way I live is inconsistent with my belief. Consistency requires a harmony between our ethical beliefs and how we live.
0068

Many consistency arguments in ethics depend on the *universalizability principle*, on which nearly all philosophers agree. Universalizability claims that whatever is right (wrong, good, bad, etc.) in one case also would be right (wrong, good, bad, etc.) in any exactly or relevantly similar case, regardless of the individuals involved. Here's an example adapted from the Good Samaritan parable (Luke 10:30–5). Suppose that, while I'm jogging, I see a man who's been beaten, robbed, and left to die. Should I help him, perhaps by making a phone call? I think of excuses why I shouldn't. I'm busy, don't want to get involved, and so on. I say to myself, "It would be all right for me not to help him." But then I consider an exactly reversed situation. I imagine myself in his place; I'm the one who's been beaten, robbed, and left to die. And I imagine him being in my place; he's jogging, sees me in my sad state, and has the same excuses. I ask myself, "Would it be all right for this man not to help me in this situation? Surely not!" But then I'm inconsistent. What's all right for me to do to another has to be all right for the other to do to me in an imagined exactly reversed situation.[1]

4.3a Exercise

Construct a turnaround argument based on the three incompatible statements in the box. Include statement C as a premise of your argument.

A. There are no universal duties.

B. Everyone ought to respect the dignity of others.

C. If everyone ought to respect the dignity of others, then there are universal duties.

[1] For more on consistency in ethics, see Chapters 13 and 14 of this present book and Chapters 7 to 9 of my *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018).

If everyone ought to respect the dignity of others, then there are universal duties.
Everyone ought to respect the dignity of others.
∴ There are universal duties.

1. Construct a different turnaround argument based on the three statements in this first box. Again, include statement C as a premise of your argument.

2. Construct a turnaround argument based on the four incompatible statements in this second box. Include statement A as a premise of your argument.

A. If we have ethical knowledge, then either ethical truths are provable or there are self-evident ethical truths.

B. We have ethical knowledge.

C. Ethical truths aren't provable.

D. There are no self-evident ethical truths.

3. Following the directions in 2, construct a second such turnaround argument.

4. Following the directions in 2, construct a third such turnaround argument. 0069

5. Construct a turnaround argument based on the three incompatible statements in this third box.

A. All human concepts come from sense experience.

B. The concept of logical validity is a human concept.

C. The concept of logical validity doesn't come from sense experience.

6. Following the directions in 5, construct a second such turnaround argument.

7. Following the directions in 5, construct a third such turnaround argument.

8. If an argument is valid, then is its turnaround necessarily also valid? Argue for the correctness of your answer.

The next seven examples are self-refuting statements. Explain how each self-refutes.

9. No statement is true.

10. Every rule has an exception.

11. One ought not to accept statements that haven't been proved.

12. Any statement whose truth or falsity we can't decide through scientific experiments is meaningless.

13. There's no such thing as something being "true." There are only opinions, each being "true for" the person holding it, none being just "true."

14. We can know only what's been proved using experimental science. I know this.

15. It's impossible to express truth in human concepts.

4.4 Constructing arguments

This book presents many logical tools; these can help turn mushy thinking into clear reasoning. You should use these logical tools where appropriate in your own reading and writing.

Imagine that your ethics teacher gives you this assignment:

> Suppose you work for a small company called Mushy Software. You can get a big contract for your company, but only by bribing an official of Enormity Incorporated. Would it be right for you to offer the bribe? Write a paper taking a position on this. Give a clear argument explaining the reasoning behind your answer.

From your study of logic, you know that an argument is a set of statements divided into premises and a conclusion. The assignment tells you to construct a valid argument along these lines:

[Insert plausible premise.]
[Insert plausible premise.]
∴ Offering the bribe is / isn't right.

Phrase your argument as clearly and simply as possible, and make sure that it's 0070 valid in some acceptable logical system. After sketching various arguments, you might arrive at this (which is valid in syllogistic and quantificational logic):

No dishonest act is right.
Offering the bribe is a dishonest act.
∴ Offering the bribe isn't right.

When you propose an argument, it's wise to ask how an opponent could object to it. While the form here is clearly valid, there might be some difficulty with the premises. How could an opponent attack the premises?

One way to attack a universal claim is to find a counterexample:

> **Counterexample:** To refute "all A is B," find something that's A but not B; to refute "no A is B," find something that's A and also B.

Premise 1 says "No dishonest act is right." You could refute this by finding an action that's dishonest and also right. Can you think of any such action? Imagine a case in which the only way to provide food for your starving family is by stealing. Presumably, stealing here is dishonest but also right:

This act of stealing is a dishonest act.
This act of stealing is right.
∴ Some dishonest acts are right.

This is valid in syllogistic and quantificational logic. So if the premises here are true, then premise 1 of your original argument is false.

Modus tollens gives another simple way to attack a claim:

> ***Modus tollens***
>
> To refute claim A, find a clearly false claim B that A implies. Then argue as below:
>
> If A then B.
> Not-B.
> ∴ Not-A.

Try to find some clearly false claim that one of the premises implies. This argument seems to work:

If no dishonest act is right, then it wouldn't be right to steal food for your starving family when this is needed to keep them from starving.
It would be right to steal food for your starving family when this is needed to keep them from starving.
∴ Some dishonest acts are right. 0071

This is valid in propositional logic. If the premises are true, then premise 1 of your original argument is false. This *modus tollens* objection is similar in intent to the counterexample objection, but phrased differently.

How can you respond to the objection? You have three options:

- *Counterattack*: Attack the arguments against your premise.
- *Reformulate*: Reword your original premises so they avoid the objection but still lead to your conclusion.
- *Change strategy*: Trash your argument and try another approach.

On the *counterattack* option, you'd maintain that the arguments against your premise either are invalid or else have false premises. Here you might claim that stealing is wrong in this hypothetical case. This would be *biting the bullet* – taking a stand that seems to go against common sense in order to defend your theory. Here you'd claim that it's wrong to steal to keep your family from starving; this is a difficult bullet to bite.

On the *reformulate* option, you'd rephrase premise 1 to avoid the objection but still lead to your conclusion. You might add the italicized qualification: "No dishonest act *that isn't needed to avoid disaster* is right." You'd have to explain what "avoid disaster" here means and you'd have to add another premise that says "Offering the bribe isn't needed to avoid disaster." Then you'd look for further objections to the revised argument.

On the *change strategy* option, you'd trash your original argument and try another approach. You might, for example, argue that offering the bribe is right (or wrong) because it's legal (or illegal), or accords with (or violates) the self-interest of the agent, or maximizes (or doesn't) the long-term interests of everyone affected by the action. Then, again, you'd have to ask whether there are objections to your new argument.

As you refine your reasoning, it's helpful to imagine a little debate going on. First present your argument to yourself. Then pretend to be your opponent and try to attack the argument. You might even enlist your friends to come up with objections; that's what professional philosophers do. Then imagine yourself trying to reply to your opponent. Then pretend to be your opponent and try to attack your reply. Repeat the process until you're content with the position you're defending and the argumentation behind it.

4.4a Exercise

Give a valid argument with plausible premises for or against these statements. For this exercise, you needn't believe these premises, but you have to regard them as plausible. Don't forget what you learned in Chapter 3 ("Meaning and Definitions") about the need to understand what a statement means before you defend or attack it.
0072

Any act is right if and only if it's in the agent's self-interest. (*ethical egoism*)
If ethical egoism is true, then it would be right for Jones to torture and kill you if this were in Jones's self-interest. It wouldn't be right for Jones to torture and kill you if this were in Jones's self-interest. ∴ Ethical egoism isn't true.

1. Offering the bribe is in the agent's self-interest.
2. Every act is right if and only if it's legal.
3. All acts that maximize good consequences are right.

4. Offering the bribe maximizes the long-term interests of everyone concerned.

5. Offering the bribe is a dishonest act.

6. Some wrong actions are errors made in good faith.

7. No error made in good faith is blameworthy.

8. All socially useful acts are right.

9. No acts of punishing the innocent are right.

10. The belief that there is a God is unnecessary to explain our experience.

11. All beliefs unnecessary to explain our experience ought to be rejected.

12. All beliefs that give practical life benefits are pragmatically justifiable.

13. The idea of a perfect circle is a human concept.

14. The idea of a perfect circle doesn't derive from sense experience.

15. All ideas gained in our earthly existence derive from sense experience.

[I took many examples from §2.3a. The English arguments in this book are a rich source of further problems for this exercise.]

4.5 Analyzing arguments

To get better at analyzing arguments, get into the habit of sketching a premise–conclusion version of arguments that you read or hear. Often the arguments will be as simple as a *modus tollens* ("If A then B, not-B, therefore not A"); but sometimes they'll get more complicated. It's important to listen and read carefully, with the aim of getting at the heart of the reasoning.

Here are four steps that you may find helpful in analyzing arguments in things you read. The steps are especially useful when you write an essay on an author's reasoning; but you also can use them to critique your own writing. The steps assume that the passage contains reasoning (and not just description).

1. *Formulate the argument in English.* Identify and write out the premises and conclusion. Aim for a valid argument expressed clearly and directly. Use the *principle of charity*: interpret unclear reasoning in the way that gives the best argument. 0073

2. Supply implicit premises where needed, avoid emotional terms, and phrase similar ideas in similar words. This step can be difficult if the author's argument is unclear.

3. *Translate into some logical system and test for validity.* If the argument is invalid, you might try step 1 again with a different formulation. If you can't get a valid argument, you can skip the next two steps.

4. *Identify difficulties.* Star controversial premises. Underline obscure or ambiguous terms; explain what you think the author means by these.

5. *Appraise the premises.* Try to decide if the premises are true. Look for informal fallacies, especially circularity and ambiguity. Give further arguments (your own or the author's) for or against the premises.

Let's try this on a famous passage from David Hume:

> Since morals, therefore, have an influence on the actions and affections, it follows, that they cannot be deriv'd from reason; and that because reason alone, as we have already prov'd, can never have any such influence. Morals excite passions, and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality, therefore, are not conclusions of our reason. No one, I believe, will deny the justness of this inference; nor is there any other means of evading it, than by denying that principle, on which it is founded. As long as it is allow'd, that reason has no influence on our passions and actions, 'tis in vain to pretend, that morality is discover'd only by a deduction of reason. An active principle can never be founded on an inactive ....[1]

First read the passage several times. Focus on the reasoning and try to put it into words; it usually takes several tries to get a clear argument. Here our analysis might look like this:

All moral judgments influence our actions and feelings.
Nothing from reason influences our actions and feelings.
∴ No moral judgments are from reason.

Next translate into some logical system and test for validity. Here we could use either syllogistic or quantificational logic:

all M is I
no R is I
∴ no M is R

$(x)(Mx \supset Ix)$
$\sim(\exists x)(Rx \bullet Ix)$
$\therefore \sim(\exists x)(Mx \bullet Rx)$

The argument tests out valid in either case.

Next identify difficulties. Star controversial premises and underline obscure or ambiguous terms:

[1] David Hume, *A Treatise of Human Nature* (Oxford: Clarendon Press, 1888), page 457 (Book III, Part I, Section I).

* All moral judgments **<u>influence</u>** our actions and feelings.
* Nothing from **<u>reason</u>** **<u>influences</u>** our actions and feelings.
∴ No moral judgments are from **<u>reason</u>**. 0074

Try to figure out what Hume meant by these underlined words. By "reason," Hume seems to mean "the discovery of truth or falsehood." Thus we can rephrase his argument as follows:

* All moral judgments **<u>influence</u>** our actions and feelings.
* No discovery of truth or falsehood **<u>influences</u>** our actions and feelings.
∴ No moral judgments are a discovery of truth or falsehood.

"Influences" also is tricky. "X influences Y" could have either of two meanings:

"X *independently of our desires* influences Y."
"X *when combined with our desires* influences Y."

Finally, appraise the premises. If we take "influences" in the first sense, then there's a problem with premise 1, which would then mean "All moral judgments, *independently of our desires*, influence our actions and feelings." This seems false, since there are people who accept moral judgments but have no desire or motivation to follow them; the actions and feelings of such a person thus wouldn't be influenced by these moral judgments. If we take "influences" in the second sense, then there's a problem with premise 2, which would then mean "No discovery of truth or falsehood, *when combined with our desires*, influences our actions and feelings." This also seems false, since the discovery of the truth that this flame would burn our finger, combined with our desire not to get burned, surely influences our actions and desires. Hume's argument is plausible because "influences" is ambiguous. Depending on how we take this term, one premise or the other becomes false or doubtful. So Hume's argument is flawed.

Here we've combined formal techniques (expressing an argument in a logical system) with informal methods (common-sense judgments, definitions, and the fallacy of ambiguity). We've used these to formulate and criticize an argument on the foundations of ethics. Our criticisms, of course, might not be final. A Hume defender might attack our arguments against Hume's premises, suggest another reading of the argument, or rephrase the premises to avoid our criticisms. But our criticisms, if clearly and logically expressed, will likely move the discussion forward. At its best, philosophical discussion involves reasoning together in a clear-headed, logical manner.

It's important to be fair when we criticize another's reasoning. Such criticism can be part of a common search for truth; we shouldn't let it descend into a vain attempt to score points. In appraising the reasoning of others, we should follow the same standards of fairness that we want others to follow in their appraisal of *our* reasoning. Distortions and other fallacies are beneath the dignity of beings, such as ourselves, who are capable of reasoning.

5 Inductive Reasoning

Much of our reasoning deals with probabilities. We observe patterns and conclude that, based on these, such and such a belief is *probably* true. This is inductive reasoning.

5.1 The statistical syllogism

The Appalachian Trail (AT), a 2,160-mile footpath from Georgia to Maine in the eastern US, has a series of lean-to shelters. Suppose we backpack on the AT and plan to spend the night at Rocky Gap Shelter. We'd like to know beforehand whether there's water (a spring or stream) close by. If we knew that *all* AT shelters have water, or that *none* do, we could reason *deductively*:

All AT shelters have water.
Rocky Gap is an AT shelter.
∴ Rocky Gap has water.

No AT shelters have water.
Rocky Gap is an AT shelter.
∴ Rocky Gap doesn't have water.

Both are deductively valid. Both have a tight connection between premises and conclusion; if the premises are true, the conclusion *has* to be true. Deductive validity is "all or nothing." Deductive arguments can't be "half-valid," nor can one be "more valid" than another.

In fact, most of the shelters have water, but a few don't. Of the shelters that I've visited, roughly 90 percent (depending on season and rainfall) have had water. If we knew that 90 percent had water, we could reason *inductively*:

90 percent of AT shelters have water.
Rocky Gap is an AT shelter.
That's all we know about the matter.
∴ Probably Rocky Gap has water.

This is a strong inductive argument. Relative to the premises, the conclusion is a good bet. But it could turn out false, even though the premises are all true.

The "That's all we know about the matter" premise means "We have no further information that influences the conclusion's probability." Suppose we just met a thirsty backpacker complaining that the water at Rocky Gap had dried up; 0076 that would change the conclusion's probability. The premise claims that we have no such further information.

Inductive arguments differ from deductive ones in two ways. (1) Inductive arguments vary in how strongly the premises support the conclusion; "99 percent of AT shelters have water" supports the conclusion more strongly than does "60 percent of AT shelters have water." We have shades of gray here – not the black and white of deductive validity/invalidity. (2) Even a strong inductive argument has only a loose connection between premises and conclusion. The premises make the conclusion at most only highly probable; the premises might be true while the conclusion is false. Inductive reasoning is a form of guessing based on recognizing and extending known patterns and resemblances.

So a **deductive argument** claims that it's *logically necessary* that if the premises are all true, then so is the conclusion. An **inductive argument** claims that it's *likely* (but not logically necessary) that if the premises are all true, then so is the conclusion. This chapter focuses on inductive arguments.

If we refine our conclusion to specify a numerical probability, we get the classic *statistical syllogism* form:

Statistical Syllogism

N percent of A's are B's.
X is an A.
That's all we know about the matter.
∴ It's N percent probable that X is a B.

90 percent of AT shelters have water.
Rocky Gap is an AT shelter.
That's all we know about the matter.
∴ *It's 90 percent probable that* Rocky Gap has water.

Here's another example:

50 percent of coin tosses are heads.
This is a coin toss.
That's all we know about the matter.
∴ It's 50 percent probable that this is heads.

Suppose that all we know affecting the probability of the toss being heads is that 50 percent of coin tosses are heads and that this is a coin toss. Then it's 50 percent probable to us that the toss is heads. This holds if we hadn't yet tossed the coin, or if we tossed it but didn't yet know how it landed. The matter is different if we know how it landed. Then it's no longer just 50 percent probable to us that it's heads; rather, we *know* that it's heads or that it's tails.

Statistical syllogisms apply most cleanly if we know little about the subject. Suppose we know these two facts about Michigan's football team:

- Michigan has first down and runs 70 percent of the time on first down.
- Michigan is behind and passes 70 percent of the time when it's behind. 0077

Relative to the first fact, Michigan probably will run. Relative to the second fact, Michigan probably will pass. But it's unclear what Michigan probably will do relative to both facts. It gets worse if we add facts about the score, the time left, and the offensive formation. Each fact by itself may lead to a clear conclusion about what Michigan probably will do; but the combination muddies the issue. Too much information can confuse us when we apply statistical syllogisms.

Chapter 1 distinguished *valid* from *sound* deductive arguments. *Valid* asserts a correct relation between premises and conclusion, but says nothing about the truth of the premises; *sound* includes both "valid" and "has true premises." It's convenient to have similar terms for inductive arguments. Let's say that an argument is *strong* inductively if the conclusion is probable relative to the premises. And let's say that an argument is *reliable* inductively if it's strong and has true premises. So then:

- With DEDUCTIVE ARGUMENTS: a correct premise–conclusion link makes the argument VALID; and VALID plus true premises makes the argument SOUND.
- With INDUCTIVE ARGUMENTS: a correct premise–conclusion link makes the argument STRONG; and STRONG plus true premises makes the argument RELIABLE.

Here's a very strong inductive argument that isn't reliable:

Michigan loses 99 percent of the times it plays.
Michigan is playing today.
That's all we know about the matter.
∴ Probably Michigan will lose today.

This is very strong, because relative to the premises the conclusion is very probable. But the argument isn't reliable, since premise 1 is false.

5.2 Probability calculations

Sometimes we can calculate probabilities precisely. Coins tend to land heads half the time and tails the other half; so each coin has a 50 percent chance of landing heads and a 50 percent chance of landing tails. Suppose we toss two coins. There are four possible combinations of heads (**H**) and tails (**T**) for the two coins:

HH HT TH TT

Each case is equally probable. So our chance of getting two heads is 25 percent (.25 or ¼), since it happens in 1 out of 4 cases. Here's the rule (where "prob" is short for "the probability" and "favorable cases" are those in which A is true): 0078

This rule holds if every case is equally likely:		
Prob of A	=	the number of favorable cases / the total number of cases

Our chance of getting at least one head is 75 percent (.75 or ¾), since it happens in 3 of 4 cases.

With odds, the ratio concerns favorable and unfavorable cases ("unfavorable cases" are those in which A is false). The odds are in your favor if the number of favorable cases is greater (then your probability is greater than 50 percent):

The odds *in favor of* A	=	the number of favorable cases / the number of unfavorable cases

So the odds are 3 to 1 in favor of getting at least one head – since it happens in 3 cases and fails in only 1 case. The odds are against you if the number of unfavorable cases is greater (so your probability is less than 50 percent):

The odds *against* A = the number of unfavorable cases / the number of favorable cases

Odds are usually given in whole numbers, with the larger number first. We wouldn't say "The odds are 1 to 3 *in favor of* getting two heads"; rather, we'd put the larger number first and say "The odds are 3 to 1 *against* getting two heads." Here are examples of how to convert between odds and probability:

- The odds are even (1 to 1) that we'll win = The probability of our winning is 50 percent.
- The odds are 7 to 5 in favor of our winning = The probability of our winning is 7/12 (7 favorable cases out of 12 total cases, or 58.3 percent).
- The odds are 7 to 5 against our winning = The probability of our winning is 5/12 (5 favorable cases out of 12 total cases, 41.7 percent).
- The probability of our winning is 70 percent = The odds are 7 to 3 in favor of our winning (70 percent favorable to 30 percent unfavorable).
- The probability of our winning is 30 percent = The odds are 7 to 3 against our winning (70 percent unfavorable to 30 percent favorable).

We'll now learn some rules for calculating probabilities. The first two rules are about necessary truths and self-contradictions: 0079

If A is a necessary truth: Prob of A = 100 percent.

If A is a self-contradiction: Prob of A = 0 percent.

Our chance of a specific coin being *either heads or not heads* is 100 percent. And our chance of it being *both heads and not heads* (at one time) is 0 percent.

This next rule relates the probability of a given event happening to the probability of that event not happening:

Prob of not-A = 100 percent – prob of A.

So if our chance of getting two heads is 25 percent, then our chance of *not* getting two heads is 75 percent (100 percent – 25 percent).

The next rule concerns events that are independent of each other, in that the occurrence of one doesn't make the occurrence of the other any more or any less likely (the first coin being heads, for example, doesn't make it any more or any less likely that the second coin will be heads):

> If A and B are independent:
>
> Prob of (A and B) = prob of A • prob of B.

Probabilities multiply with AND. So our chance of throwing two heads (25 percent) *and then* throwing two heads again (25 percent) is 6.25 percent (25 percent • 25 percent).

This next rule holds for events that are mutually exclusive, in that they can't both happen together:

> If A and B are mutually exclusive:
>
> Prob of (A or B) = prob of A + prob of B.

Probabilities add with OR. It can't happen that we throw two heads and also (on the same toss of two coins) throw two tails. The probability of either event is 25 percent. So the probability of one *or* the other happening (getting two heads *or* two tails) is 50 percent (25 percent + 25 percent). When the two events aren't mutually exclusive, we use this more complex rule:

> This holds even if A and B aren't mutually exclusive:
>
> Prob of (A or B) = Prob of A + prob of B – prob of (A and B).

Suppose we calculate the probability of getting at least one head when we flip
0080 two coins. Coin 1 being heads and coin 2 being heads aren't mutually exclusive, since they might both happen together; so we apply the more complex rule. The chance of coin 1 being heads *or* coin 2 being heads = the chance of coin 1 being heads (50 percent) + the chance of coin 2 being heads (50 percent) – the chance of coin 1 and coin 2 both being heads (25 percent). So our chance of getting at least one head is 75 percent (50 + 50 – 25). If A and B are mutually exclusive, then the probability of (A and B) = 0 and the simpler rule gives the same result.

Suppose we throw two dice. There are six equally probable possibilities for each die. Here are the possible combinations and resulting totals (the numbers on the left are for the first die, the numbers on the top are for the second die, and the other numbers are the totals):

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

These 36 combinations each have an equal 1/36 probability. The chance of getting 12 is 1/36, since we get 12 in only 1 of 36 cases. The chance of getting 11 is 1/18 (2/36) – since we get 11 in 2 of 36 cases. Similarly, we have a 1/6 (6/36) chance of getting 10 or higher, and a 5/6 (30/36) chance of getting 9 or lower.

Suppose we have a standard deck of 52 cards. What's our chance of getting 2 aces when dealt 2 cards? We might think that, since 1/13 of the cards are aces, our chance of getting two aces is 1/169 (1/13 • 1/13). But that's wrong. Our chance of getting an ace on the first draw is 1/13, since there are 4 aces in the 52 cards, and 4/52 = 1/13. But if we get an ace on the first draw, then there are only 3 aces left in the 51 cards. So our chance of getting a second ace is 1/17 (3/51). Thus, our chance of getting 2 aces is 1/221 (1/13 • 1/17), or about 0.45 percent.

Here the events aren't independent. Getting an ace on the first card reduces the number of aces left and our chance of drawing an ace for the second card. This is unlike coins, where getting heads on one toss doesn't affect our chance of getting heads on the next toss. If events A and B aren't independent, we need this rule for determining the probability of the conjunction (A and B):

> This holds even if A and B aren't independent:
>
> Prob of (A and B) = Prob of A • (prob of B after A occurs).

This reflects the reasoning about our chance of getting 2 aces from a 52-card deck. What's our chance with a double 104-card deck? Our chance of getting a first ace is again 1/13 (since there are 8 aces among the 104 cards, and 8/104 = 1/13). After we get a first ace, there are 7 aces left in the 103 cards, and so our chance of a second ace is 7/103. So the probability of getting a first ace and then 0081 a second ace = 1/13 (the probability of the first ace) • 7/103 (the probability of the second ace). This works out to 7/1339 (1/13 • 7/103), or about 0.52 percent. So our chance of getting 2 aces when dealt 2 cards from a double 104-card deck is about 0.52 percent (or slightly better than the 0.45 with a standard deck).

Mathematically fair betting odds are in reverse proportion to probability.

Suppose we bet on whether, in drawing 2 cards from a standard 52-card deck, we'll draw 2 aces. There's a 1/221 chance of this, so the odds are 220 to 1 against us. If we bet $1, we should get $220 if we win. If we play for a long time under such betting odds, our gains and losses will likely roughly equalize. In a casino, the house takes its cut and so we get a lower payoff. So if we play there a long time under such odds, probably we'll lose and the casino will win. That's why Las Vegas casinos look like the palaces of emperors.

5.2a Exercise: LogiCola P (P, O, & C)

Work out the following problems. A calculator is useful for some of them.

You're playing blackjack and your first card is an ace. What's your chance of getting a card worth 10 (a 10, jack, queen, or king) for your next card? You're using a standard 52-card deck.
There are 16 such cards (one 10, J, Q, and K for each suit) from 51 remaining cards. So your chance is 16/51 (about 31.4 percent).

1. What would the answer to the sample problem be with a double 104-card deck?

2. Suppose the Cubs and Mets play baseball today. There's a 60 percent chance of rain, which would cancel the game. If the teams play, the Cubs have a 20 percent chance of winning. What chance do the Cubs have of winning today?

3. You're tossing coins. You tossed 5 heads in a row using a fair coin. What's the probability now that the next coin will be heads?

4. You're about to toss 6 coins. What's the probability that all 6 will be heads?

5. Suppose there's an 80 percent chance that the winner of the Michigan versus Ohio State game will go to the Rose Bowl, a 60 percent chance that Michigan will beat Ohio State, and a 30 percent chance that Michigan will win the Rose Bowl if it goes. Then what's the probability that Michigan will win the Rose Bowl?

6. Suppose you bet $10 that Michigan will win the Rose Bowl. Assuming the probabilities of the last example and mathematically fair betting odds, how much money should you win if Michigan wins the Rose Bowl?

7. You're playing blackjack and get an ace for the first card. You know that the cards used on the only previous hand were a 5, a 6, two 7's, and two 9's, and that all these are in the discard pile. What's your chance of getting a card worth 10 (a 10, jack, queen, or king) for the next card? You're using a standard 52-card deck.

8. What would the answer to the last problem be with a double 104-card deck?

9. You're throwing a pair of dice. Your sister bets you even money that you'll throw an even number (adding both together). Is she playing you for a sucker?

10. Your sister is throwing a pair of dice. She says, “I bet I’ll throw a number divisible by three.” What are the mathematically fair betting odds? 0082

11. You’re dealt five cards: two 3s, a 4, a 6, and a 7. If you get another card, what’s the probability that it will be a 5? What’s the probability that it will be a 3?

12. You’re at a casino in Las Vegas and walk by a $1 slot machine that says “Win $2,000!” Assume that this is the only way you can win and that it gives mathematically fair odds or worse. What’s your chance of winning if you deposit $1?

13. What’s the probability, ignoring leap-year complications, that both your parents have their birthday on the same day of the year (whatever day that may be)?

14. Our football team, Michigan, is 2 points behind with a few seconds left. We have the ball, fourth and two, on the Ohio State 38. We could have the kicker try a long field goal, which would win the game. The probability of kicking this goal is 30 percent. Or we could try to make a first down and then kick from a shorter distance. There’s a 70 percent probability of making a first down and a 50 percent probability of making the shorter field goal if we make the first down. Which alternative gives us a better chance to make the field goal?

15. Our team, Michigan, is 2 points ahead with a minute left. Ohio State is going for it on fourth down. It’s 60 percent probable that they’ll pass, and 40 percent probable that they’ll run. We can defense the pass or defense the run. If we defense the pass, then we’re 70 percent likely to stop a pass but only 40 percent likely to stop a run. If we defense the run, then we’re 80 percent likely to stop a run but only 50 percent likely to stop a pass. What should we do?

5.3 Philosophical questions

We’ll now consider four philosophical questions on probability. Philosophers disagree about how to answer these questions.

1. Are the ultimate scientific laws governing the universe *deterministic* or *probabilistic* in nature?

Some think all ultimate scientific laws are *deterministic*; we use probability only because we lack knowledge. Suppose we knew all scientific laws and the complete state of the world at a given time. Then we could in principle infallibly predict whether the coin will come up heads, whether it will rain three years from today, and who will win the World Cup in 30 years. This is the thesis of determinism.

Others say that some or all of the ultimate laws governing our world are *probabilistic*. Such laws say that under given conditions a result will *probably* obtain, but not that it *must* obtain. The world is a dice game.

The empirical evidence on this issue is inconclusive. Quantum physics today embraces probabilistic laws but could someday return to deterministic laws. The issue is complicated by the controversy over whether determinism is an empirical or an *a priori* issue (§3.7); some think reason (not experience) gives us certainty that the world is deterministic. 0083

2. What does "probable" mean? And can every statement be assigned a numerical probability relative to given evidence?

"Probable" has various senses. "The *probability* of heads is 50 percent" could be taken in at least four ways:

- *Ratio of observed frequencies*: We've observed that coins land heads about half of the time.
- *Ratio of abstract possibilities*: Heads is one of the two equally likely abstract possibilities.
- *Measure of actual confidence*: We have the same confidence in the toss being heads as we have in it not being heads.
- *Measure of rational confidence*: It's rational to have the same confidence in the toss being heads as in it not being heads.

We used a *ratio of observed frequencies* to calculate the probability of finding water at Rocky Gap Shelter. And we used a *ratio of abstract possibilities* to calculate the probability of being dealt two aces. But sometimes these ratio approaches can't give numerical probabilities. Neither ratio approach gives a numerical probability to "Michigan will run" relative to information about ancient Greek philosophy or relative to this combination:

- Michigan has first down and runs 70 percent of the time on first down.
- Michigan is behind and passes 70 percent of the time when it's behind.[1]

Only in special cases do the ratio approaches give numerical probabilities.

The *measure of actual confidence* sometimes yields numerical probabilities. Consider these statements:

"There's life on other galaxies."
"Michigan will beat Ohio State this year."
"There's a God."

If you regard 1-to-1 betting odds on one of these as fair, then your actual

[1] Here it would be helpful to know what Michigan does on first down when they're behind. But the same problem continues if other factors are relevant (e.g., how much time is left in the game).

confidence in the statement is 50 percent. But you may be unwilling to commit yourself to such odds. Maybe you can't say if your confidence in the statement is less or greater than your confidence that a coin toss will be heads. Then we can't assign numbers to your actual confidence. The *rational confidence view*, too, would have trouble assigning numerical probabilities in these cases.

Some doubt that probability as rational confidence satisfies the standard probability rules of the last section. These rules say that necessary statements always are 100 percent probable. But consider a complex propositional logic formula 0084 that's a necessary truth, even though your evidence suggests that it isn't; perhaps your normally reliable logic teacher tells you that it's not a necessary truth – or perhaps in error you get a truth-table line of false (see §6.5). Relative to your data, it seems rational not to put 100 percent confidence in the formula, even though it in fact is a necessary truth. So is probability theory wrong?

Probability theory is idealized rather than wrong. It describes the confidence an ideal reasoner would have; an ideal reasoner would always recognize necessary truths and put 100 percent confidence in them. So we have to be careful applying probability theory to the beliefs of non-ideal human beings; we must be like physicists who give simple equations for frictionless bodies and then keep in mind that these are idealized when applying the equations to real cases.

Probability as *actual confidence* definitely can violate the probability rules. Many would calculate the probability of drawing 2 aces from a 52 or 104 card deck as 1/169 (1/13 • 1/13); so they'd regard 168-to-1 betting odds as fair. But the probability rules say this is wrong (§5.2).

3. How does probability relate to how ideally rational persons *believe*?

Some think ideally rational persons would believe all and only those statements that are more than 50 percent probable. But this has strange implications. Suppose that Austria, Brazil, and China each has a $33\frac{1}{3}$ percent chance of winning the World Cup. Then each of these is $66\frac{2}{3}$ percent probable:

"Austria won't win the World Cup, but Brazil or China will."
"Brazil won't win the World Cup, but Austria or China will."
"China won't win the World Cup, but Austria or Brazil will."

On the view just described, ideally rational persons would believe all three statements. But this is silly; only a very confused person could do this.

The view has other problems. Why pick 50 percent? Why wouldn't ideally rational persons believe all and only those statements that are at least 60 percent (or 75 percent or 90 percent) probable? And there are further problems if there's no way to work out numerical probabilities.

The view gives an ideal of selecting all beliefs in a way that's free of subjective factors (like feelings and practical interests). Some find this ideal attractive. Pragmatists find it repulsive. They believe in following subjective factors on issues that our intellects can't decide. They think that numerical probability doesn't apply to life's deeper issues (like free will, God, or basic moral principles).

4. How does probability relate to how ideally rational persons *act*?

Some think ideally rational persons always act to maximize **expected gain**. In working out what to do, they'd list the possible alternative actions (A, B, C, …) and then consider the possible outcomes (A1, A2, A3, …) of each action. The gain or loss of each outcome 0085 would be multiplied by the probability of that outcome occurring; adding these together gives the action's *expected gain*. So an action's expected gain is the sum of probability-times-gain of its various possible outcomes. Ideally rational persons, on this view, would always do whatever had the highest expected gain (or the lowest expected loss when all alternatives lose).

What is "gain" here? Is it pleasure or desire-satisfaction, for oneself or one's group or all affected by the action? Or is it financial gain, for oneself or one's company? Consider an economic version of the theory, that ideally rational gamblers would always act to maximize their expected financial gain. Imagine that you're such an "ideally rational gambler." You find a game of dice that pays $3,536 on a $100 bet if you throw 12. You'd work out the expected gain of playing or not playing (alternatives P and N) in this way:

P. PLAYING. There are two possible outcomes: P1 (I win) and P2 (I lose). P1 is 1/36 likely and gains $3,536; P1 is worth (1/36 • $3,536) or $98.22. P2 is 35/36 likely and loses $100; P2 is worth (35/36 • –$100), or –$97.22. The expected gain of alternative P is ($98.22 – 97.22), or $1.

N. NOT PLAYING. On this alternative, I won't win or lose anything. The expected gain of alternative N is (100 percent • $0), or $0.

So then you'd play. If you played this dice game only once, you'd be 97 percent likely to lose money. But the occasional payoff is great; you'd likely gain about a million dollars if you played a million times.

"Ideally rational gamblers" would gamble if the payoff were favorable, but not otherwise. Since casinos take their cut, their payoff is lower; ideally rational gamblers wouldn't gamble there. But people have interests other than money; for many, gambling is great fun, and they're willing to pay for the fun.

Some whose only concern is money refuse to gamble even when the odds are in their favor. Their concern may be to have *enough* money. They may better satisfy this by being cautious; they don't want to risk losing what they have for the sake of gaining more. Few people would endanger their total savings for the 1-in-900 chance of gaining a fortune 1000 times as great.

Another problem with the "maximize expected gain" policy is that it's often difficult or impossible to give objective numerical probabilities and to multiply probability by gain. So this policy faces grave difficulties if taken as an overall guide to life. But it can sometimes be useful as a rough guide. At times it's helpful to work out the expected gain of the various alternatives, perhaps guessing at the probabilities and gains involved.

I once had two alternatives in choosing a flight:

Ticket A costs \$250 and allows me to change my return date.
Ticket B costs \$200 and has a \$125 charge if I change my return date.

Which ticket is a better deal for me? Intuitively, A is better if a change is very likely, while B is better if a change is very unlikely. But we can be more precise. 0086 Let x represent the probability of my changing the return. Then:

Expected cost of A = \$250.
Expected cost of B = \$200 + (\$125 • x).

Algebra shows the expected costs are identical if x is 40 percent. So A is better if a change is more than 40 percent likely, while B is better if a change is less likely than that. Judging from past experiences, the probability of my changing the return date was less than 40 percent. Thus, ticket B minimized my expected cost. So I bought ticket B.

In some cases, however, it might be more rational to pick A. Maybe I have \$250 but I don't have the \$325 that option B might cost me; so I'd be in great trouble if I had to change the return date. It might then be more rational to follow "better safe than sorry" and pick A.

5.3a Exercise: LogiCola P (G, D, & V)

Suppose you decide to believe all and only statements that are more probable than not. You're tossing three coins; which of the next six statements would you believe?

Either the first coin will be heads, or all three will be tails.
You'd believe this, since it happens in 5 of 8 cases: **HHH HHT HTH HTT** THH THT TTH **TTT**

1. I'll get three heads.

2. I'll get at least one tail.

3. I'll get two heads and one tail.

4. I'll get either two heads and one tail, or else two tails and one head.

5. The first coin will be heads.

For problems 6 through 10, suppose you decide to do in all cases whatever would maximize your expected financial gain.

6. You're deciding whether to keep your life savings in a bank (which pays a dependable 10 percent) or invest in Mushy Software. If you invest in Mushy, you have a 99 percent chance of losing everything and a 1 percent chance of making 120 times your investment this year. What should you do?

7. You're deciding whether to get hospitalization insurance. There's a 1 percent chance per year that you'll have a $10,000 hospital visit (ignore other hospitalizations); the insurance would cover it all. What's the most you'd agree to pay per year for this insurance?

8. You're running a company that offers hospitalization insurance. There's a 1 percent chance per year that a customer will have a $10,000 hospital visit (ignore other hospitalizations); the insurance would cover it all. What's the least you could charge per year for this insurance to likely break even? 0087

9. You're deciding whether to invest in Mushy Software or Enormity Incorporated. Mushy stock has a 30 percent probability of gaining 80 percent, and a 70 percent probability of losing 20 percent. Enormity stock has a 100 percent probability of gaining 11 percent. Which should you invest in?

10. You're deciding whether to buy a computer from Cut-Rate or Enormity. Both models perform identically. There's a 60 percent probability that either machine will need repair over the period you'll keep it. The Cut-Rate model is $600 but will be a total loss (requiring the purchase of another computer for $600) if it ever needs repair. The Enormity Incorporated model is $900 but offers free repairs. Which should you buy?

5.4 Reasoning from a sample

Recall our statistical syllogism about the Appalachian Trail:

90 percent of the AT shelters have water.
Rocky Gap is an AT shelter.
That's all we know about the matter.
∴ Probably Rocky Gap has water.

Premise 1 says 90 percent of the shelters have water. I might know this because I've checked all 300 shelters and found that 270 of them had water. More likely, I base my claim on inductive reasoning. On my AT hikes (and I've hiked the whole Georgia-to-Maine distance), I've observed a large and varied group of shelters, and about 90 percent have had water. I conclude that probably roughly 90 percent of *all* the shelters (including those not observed) have water:

Sample-projection syllogism
N percent of examined A's are B's. A large and varied group of A's has been examined. ∴ Probably roughly N percent of all A's are B's.
90 percent of examined AT shelters have water. A large and varied group of AT shelters has been examined. ∴ Probably roughly 90 percent of all AT shelters have water.

Such reasoning assumes that a large and varied sample probably gives us a good idea of the whole. The strength of such reasoning depends on: (1) *size* of sample; (2) *variety* of sample; and (3) *cautiousness* of conclusion.

1. Other things being equal, a *larger sample* gives a stronger argument. A projection based on a small sample (ten shelters, for example) would be weak. My sample included about 150 shelters.

2. Other things being equal, a *more varied sample* gives a stronger argument. A sample is *varied* to the extent that it proportionally represents the diversity of the whole. AT shelters differ. Some are on high ridges, others are in valleys. 0088 Some are on the main trail, others are on blue-blazed side trails. Some are in wilderness areas, others are in rural areas. Our sample is varied to the extent that it reflects this diversity.

We'd have a weak argument if we examined only the dozen or so shelters in Georgia. This sample is small, has little variety, and covers only one part of the trail; but the poor sample might be all we have. Background information can help us to criticize a sample. Suppose we checked only AT shelters located on mountain tops or ridges. If we knew that water tends to be scarcer in such places, we'd judge this sample to be biased.

3. Other things being equal, we get a stronger argument if we have a *more cautious conclusion*. We have stronger reason for thinking the proportion of shelters with water is "between 80 and 95 percent" than for thinking that it's "between 89 and 91 percent." Our original argument says "roughly 90 percent." This is vague; whether it's too vague depends on our purposes.

Suppose our sample-projection argument is strong and has premises all true. Then it's likely that roughly 90 percent of the shelters have water. But the conclusion is only a rational guess; it could be far off. It's may even happen that every shelter that we didn't check is dry. Inductive reasoning brings risk.

Here's another sample-projection argument:

52 percent of the voters we checked favor the Democrat.
A large and varied group of voters has been checked.
∴ Probably roughly 52 percent of all voters favor the Democrat.

Again, our argument is stronger if we have a larger and more varied sample and a more cautious conclusion. A sample of 500 to 1000 people supposedly yields a margin of likely error of less than 5 percent; we should then construe our conclusion as "Probably between 57 percent and 47 percent of all voters favor the Democrat." To get a varied sample, we might select people using a random process that gives everyone an equal chance of being included. We also might try to have our sample proportionally represent groups (like farmers and the elderly) that tend to vote in a similar way. We should word our survey fairly and not intimidate people into giving a certain answer. And we should be clear whether we're checking registered voters or probable voters. Doing a good pre-election survey isn't easy.

A sample-projection argument ends the way a statistical syllogism begins – with "N percent of all A's are B's." It's natural to connect the two:

90 percent of examined AT shelters have water.
A large and varied group of AT shelters has been examined.
∴ Probably roughly 90 percent of all AT shelters have water.
Rocky Gap is an AT shelter.
That's all we know about the matter.
∴ It's roughly 90 percent probable that Rocky Gap has water. 0089

A sample-projection argument could use "all" instead of a percentage:

All examined cats purr.
A large and varied group of cats has been examined.
∴ Probably all cats purr.

This conclusion makes a strong claim, since a single non-purring cat would make it false; this makes the argument riskier and weaker. We could expand the argument further to draw a conclusion about a specific cat:

All examined cats purr.
A large and varied group of cats has been examined.
∴ Probably all cats purr.
Socracat is a cat.
∴ Probably Socracat purrs.

Thus sample-projection syllogisms can have various forms.

5.4a Exercise

Evaluate the following inductive arguments.

After contacting 2 million voters on telephone, we conclude that Landon will beat Roosevelt in 1936 by a landslide for the US presidency. (This was an actual prediction.)
The sample was biased. Those who could afford telephones during the Depression tended to be richer and more Republican. Roosevelt won easily.

1. I randomly examined 200 Loyola University Chicago students at the law school and found that 15 percent were born in Chicago. So probably 15 percent of all Loyola students were born in Illinois.

2. I examined every Loyola student whose Social Security number ended in 3 and I found that exactly 78.4 percent of them were born in Chicago. So probably 78.4 percent of all Loyola students were born in Chicago.

3. Italians are generally fat and lazy. How do I know? Well, when I visited Rome for a weekend last year, all the hotel employees were fat and lazy – all six of them.

4. I meet many people in my daily activities; the great majority of them intend to vote for the Democrat. So the Democrat probably will win.

5. The sun has risen every day as long as humans can remember. So the sun will likely rise tomorrow. (How can we put this into standard form?)

Consider this inductive argument: "Lucy got an A on the first four logic quizzes, so probably she'll also get an A on the fifth logic quiz." Would each of the statements 6 through 10 strengthen or weaken this argument?

6. Lucy has been sick for the last few weeks and has missed most of her classes.

7. The first four quizzes were on formal logic, while the fifth is on informal logic. 0090

8. Lucy has never received less than an A in her life.

9. A student in this course gets to drop the lowest of the five quizzes.

10. Lucy just took her Law School Admissions Test.

We'll later see a deductive version of the classic argument from design for the existence of God (§7.1b #4). The following inductive version has a sample-projection form and is very controversial. Evaluate the truth of the premises and the general inductive strength of the argument.

11. The universe is orderly (like a watch that follows complex laws).
Most orderly things we've examined have intelligent designers.
We've examined a large and varied group of orderly things.
That's all we know about the matter.
∴ The universe probably has an intelligent designer.

5.5 Analogical reasoning

Suppose you're exploring your first Las Vegas casino. The casino is huge and filled with people. There are slot machines for nickels, dimes, quarters, and dollars. There are tables for blackjack and poker. There's a big roulette wheel. There's a bar and an inexpensive all-you-can-eat buffet.

You then go into your second Las Vegas casino and notice many of the same things: the size of the casino, the crowd, the slot machines, the blackjack and poker tables, the roulette wheel, and the bar. You're hungry. Recalling what you saw in your first casino, you conclude, "I bet this place has an inexpensive all-you-can-eat buffet, just like the first casino."

This is an argument by analogy. The first and second casinos are alike in many ways, so they're probably alike in some further way:

Most things true of casino 1 also are true of casino 2.
Casino 1 has an all-you-can-eat buffet.
That's all we know about the matter.
∴ Probably casino 2 also has an all-you-can-eat buffet.

Here's a more wholesome example (about Appalachian Trail shelters):

Most things true of the first AT shelter are true of this second one.
The first AT shelter had a logbook for visitors.
That's all we know about the matter.
∴ Probably this second shelter also has a logbook for visitors.

We argue that things similar in many ways are likely similar in a further way.

Statistical and analogical arguments are closely related: 0091

Statistical

Most large casinos have buffets.
Circus Circus is a large casino.
That's all we know about the matter.
∴ Probably Circus Circus has a buffet.

Analogical

Most things true of casino 1 are true of casino 2.
Casino 1 has a buffet.
That's all we know about the matter.
∴ Probably casino 2 has a buffet.

The first rests on our experience of *many casinos*, while the second rests on our experience of *many features* that two casinos have in common.

Here's the general form of the analogy syllogism:

Analogy syllogism

Most things true of X also are true of Y.
X is A.
That's all we know about the matter.
∴ Probably Y is A.

Premise 1 is rough. In practice, we don't just count similarities; rather we look for how relevant the similarities are to the conclusion. While the two casinos were alike in many ways, they also differed in some ways:

- Casino 1 has a name whose first letter is "S," while casino 2 doesn't.
- Casino 1 has a name whose second letter is "A," while casino 2 doesn't.
- Casino 1 has quarter slot machines by the front entrance, while casino 2 has dollar slots there.

These factors aren't relevant and so don't weaken our argument that casino 2 has a buffet. But the following differences would weaken the argument:

- Casino 1 is huge, while casino 2 is small.
- Casino 1 has a bar, while casino 2 doesn't.
- Casino 1 has a big sign advertising a buffet, while casino 2 has no such sign.

These factors would make a buffet in casino 2 less likely.

So we don't just count similarities when we argue by analogy; many similarities are trivial and unimportant. Rather, we look to *relevant* similarities. But how do we decide which similarities are relevant? We somehow appeal to our background information about what things are likely to go together. It's difficult to give rules here – even vague ones.

Our "Analogy Syllogism" formulation is a rough sketch of a subtle form of reasoning. Analogical reasoning is elusive and difficult to put into strict rules.
0092

5.5a Exercise: LogiCola P (I)

Suppose you're familiar with this Gensler logic book but with no others. Your friend Sarah is taking logic and uses another book. You think to yourself, "My book discusses analogical reasoning, and so Sarah's book likely does too." Which of these bits of information would strengthen or weaken this argument – and why?

Sarah's course is a specialized graduate course on quantified modal logic.
This weakens the argument; such a course probably wouldn't discuss analogical reasoning. (This answer presumes background information.)

1. Sarah's book has a different color.
2. Sarah's book also has chapters on syllogisms, propositional logic, quantificational logic, and meaning and definitions.
3. Sarah's course is taught by a member of the math department.
4. Sarah's chapter on syllogisms doesn't use the star test.
5. Sarah's book is abstract and has few real-life examples.
6. Sarah's book isn't published by Routledge.
7. Sarah's book is entirely on informal logic.
8. Sarah's book has cartoons.
9. Sarah's book has 100 pages on inductive reasoning.
10. Sarah's book has 10 pages on inductive reasoning.

Suppose your friend Tony at another school took an ethics course that discussed utilitarianism. You're taking an ethics course next semester. You think to yourself, "Tony's course discussed utilitarianism, and so my course likely will too." Which of these bits of information would strengthen or weaken this argument – and why?

11. Tony's teacher transferred to your school and will teach your course as well.
12. Tony's course was in medical ethics, while yours is in general ethical theory.
13. Both courses use the same textbook.
14. Tony's teacher has a good reputation, while yours doesn't.
15. Your teacher is a Marxist, while Tony's isn't.

5.6 Analogy and other minds

We'll now study a classic philosophical example of analogical reasoning. This will help us to appreciate the elusive nature of such arguments.

Consider these two hypotheses:

- There are other conscious beings (with thoughts and feelings) besides me.
- I'm the only conscious being. Other humans are like cleverly constructed robots; they have outer behavior but no inner thoughts and feelings.

We all accept the first hypothesis and reject the second. How can we justify this 0093 intellectually? Consider that I can directly feel my own pain, but not the pain of others. When I experience the *pain behavior* of others, how do I know that this behavior manifests an inner experience of pain?

One approach appeals to an argument from analogy:

Most things true of me also are true of Jones. (We're both alike in general behavior, nervous system, and so on.)
I generally feel pain when showing outward pain behavior.
This is all I know about the matter.
∴ Probably Jones also feels pain when showing outward pain behavior.

Since Jones and I are alike in most respects, we're probably alike in a further respect – that we both feel pain when we show pain behavior. But then there'd be other conscious beings besides me.

Here are four ways to criticize this argument:

- Jones and I also differ in many ways. This may weaken the argument.
- Since I can't directly feel Jones's pain, I can't have direct access to the truth of the conclusion. This makes the argument peculiar and may weaken it.
- I have a sample-projection argument against there being other conscious beings: "All the conscious experiences that I've experienced are mine; but I've examined a large and varied group of conscious experiences. And so probably all conscious experiences are mine (but then I'm the only conscious being)."
- Since the analogical argument is weakened by such considerations, it at most makes it only somewhat probable that there are other conscious beings. But normally we take this belief to be solidly based.

Suppose we reject the analogical argument. Then *why* should we believe in other minds? Because it's an instinctive, commonsense belief that hasn't been disproved and that's in our practical and emotional interests to accept? Or because of a special rule of evidence, not based on analogy, that experiencing another's behavior justifies beliefs about their mental states? Or because talk about mental states is really just talk about behavior (so "being in pain" means "showing pain behavior")? Or maybe there's no answer – and I don't really *know* if there are other conscious beings besides me.

The analogical argument for other minds highlights problems with induction. Philosophers seldom dispute whether *deductive* arguments have a correct connection between premises and conclusion; instead, they dispute the truth of the premises. But with *inductive* arguments it's often disputed whether and to what extent the premises, if true, provide good reason for accepting the conclusion. Those who like things neat and tidy prefer deductive to inductive reasoning. 0094

5.7 Mill's methods

John Stuart Mill, a 19th-century British philosopher, formulated five methods for arriving at and justifying beliefs about causes. We'll study three of these. His basic idea is that factors that regularly occur together may be causally related.

Suppose that Alice, Bob, Carol, and David were at a party. Alice and David got sick, and food poisoning is suspected. Hamburgers, pie, and ice cream were served. This chart shows who ate what and who got sick (where Hm = Hamburger, Pi = Pie, IC = Ice Cream, Sk = Sick, y = yes, and n = no):

	Hm	Pi	IC	Sk
Alice	y	y	n	y
Bob	n	n	y	n
Carol	y	n	n	n
David	n	y	y	y

To find what caused the sickness, we'd search for a factor that correlates with the "yes" answers in the "sick" column. This suggests that the pie did it. Pie is the only thing eaten by all and only those who got sick. This reasoning reflects Mill's method of agreement:

Agreement

A occurred more than once.
B is the only additional factor that occurred if and only if A occurred.
∴ Probably B caused A, or A caused B.

Sickness occurred more than once.
Eating pie is the only additional factor that occurred if and only if sickness occurred.
∴ Probably eating pie caused sickness, or sickness caused the eating of pie.

The second alternative, that sickness caused the eating of pie (perhaps by bringing about a special craving?), is interesting but implausible. So we'd conclude that the people *probably* got sick because of eating the pie.

The "probably" is important. Eating the pie and getting sick might just happen to have occurred together; maybe there's no causal connection. Some unmentioned factor (maybe drinking bad water while hiking) might have caused the sicknesses. Or maybe the two sicknesses had different causes.

We took for granted a simplifying assumption. We assumed that the two cases of sickness had the same cause which was a single factor on our list and always caused sickness. Our investigation may force us to give up this assumption and consider more complex solutions. But it's good to try simple solutions first and avoid complex ones as long as we can.

We can definitely conclude that eating the hamburgers doesn't necessarily make a person sick, since Carol ate them but didn't get sick. Similarly, eating the ice cream doesn't necessarily make a person sick, since Bob ate it but didn't get sick. Let's call this sort of reasoning the "method of disagreement":
0095

Disagreement

A occurred in some case.
B didn't occur in the same case.
∴ A doesn't necessarily cause B.

Eating the ice cream occurred in Bob's case.
Sickness didn't occur in Bob's case.
∴ Eating the ice cream doesn't necessarily cause sickness.

Mill used this form of reasoning but didn't include it in his five methods.

Suppose *two* factors – eating pie and eating hamburgers – occurred in just those cases where someone got sick. Then the method of agreement wouldn't lead to any definite conclusion about which caused the sickness. To make sure it was the pie, we might do an experiment. We take two people, Eduardo and Frank, who are as alike as possible in health and diet. We give them all the same things to eat, except that we feed pie to Eduardo but not to Frank. (This is unethical, but it makes a good example.) Then we see what happens. Suppose Eduardo gets sick but Frank doesn't; then we can conclude that the pie probably caused the sickness. This follows Mill’s method of difference:

Difference

A occurred in the first case but not the second.
The cases are otherwise identical, except that B also occurred in the first case but not in the second.
∴ Probably B is (or is part of) the cause of A, or A is (or is part of) the cause of B.

Sickness occurred in Eduardo’s case but not Frank’s.
The cases are otherwise identical, except that eating pie occurred in Eduardo’s case but not Frank’s.
∴ Probably eating pie is (or is part of) the cause of the sickness, or the sickness is (or is part of) the cause of eating pie.

Since we made Eduardo eat the pie, we reject the second main alternative. So probably eating pie is (or is part of) the cause of the sickness. The cause might simply be the eating of the pie (which contained a virus). Or the cause might be this combined with one’s poor physical condition.

Another unethical experiment illustrates Mill’s method of variation. This time we find four victims (George, Henry, Isabel, and Jodi) and feed them varying amounts of pie (where 1 = tiny slice, 2 = small slice, 3 = normal slice, 4 = two slices). They get sick in varying degrees (where 1 = slightly sick, 2 = somewhat sick, 3 = very sick, 4 = wants to die):

	Pie	Sick
George	1	1
Henry	2	2
Isabel	3	3
Jodi	4	4

So the pie probably caused the sickness, following Mill’s method of variation:
0096

Variation

A changes in a certain way if and only if B also changes in a certain way.
∴ Probably B’s changes caused A’s, or A’s caused B’s, or some C caused both.

People got sicker if and only if they ate more pie.
∴ Probably eating pie caused the sickness, or the sickness caused the eating of pie, or something else caused both the eating and the sickness.

The last two alternatives are implausible. So we conclude that eating the pie probably caused the sickness.

Mill's methods often give us a conclusion with several alternatives. Temporal sequence can eliminate an alternative; the cause can't come after the effect. Suppose we conclude this: "Either laziness during previous months caused the F on the final exam, or the F on the final exam caused laziness during the previous months." Here we'd reject the second alternative.

"Cause" can mean either "total cause" or "partial cause." Suppose Jones got shot and then died. Misapplying the method of disagreement, we might conclude that being shot didn't cause the death, since some who are shot don't die. But the proper conclusion is rather that being shot doesn't necessarily cause death. We also can conclude that being shot wasn't the *total* cause of Jones's death (even though it might be a partial cause). What caused Jones's death wasn't just that he was shot. What caused the death was that he was shot in a certain way in certain circumstances (for example, through the head with no medical help). This is the total cause; anyone shot that exact way in those exact circumstances (including the same physical and mental condition) would have died. The method of disagreement deals with total causes, not partial causes.

The ambiguities of the word "cause" run deep. "Factor A causes factor B" could mean any combination of these:

> Factor A will always (or probably) by itself (or in combination with factor C) directly (or through a further causal chain) bring about factor B; or the absence of factor A will … bring about the absence of factor B; or both.

The probabilistic sense is controversial. Suppose that the incidence of lung cancer varies closely with heavy smoking, so heavy smokers are much more likely to get lung cancer. Could this probabilistic connection be enough for us to say that heavy smoking is a (partial) cause of lung cancer? Or is it wrong to use "cause" unless we have some factor C such that heavy smoking when combined with factor C *always* results in lung cancer? Part of the debate over whether a "causal connection" exists between heavy smoking and lung cancer is semantic. Can we use "cause" with probabilistic connections? If we can speak of Russian roulette *causing* death, then we can speak of heavy smoking *causing* lung cancer. 0097

5.7a Exercise: LogiCola P (M & B)

Draw whatever conclusions you can using Mill's methods; supplement Mill's methods by common sense when appropriate. Say which method you're using, what alternatives you conclude from the method itself, and how you narrow the conclusion down to a single alternative. Also say when Mill's methods lead to no definite conclusion.

Kristen's computer gave error messages when she booted up. We changed things one at a time to see what would stop the messages. What worked was updating the video driver.
By the difference method, probably updating the driver caused (or partially caused) the error messages to stop, or stopping the messages caused (or partially caused) us to update the driver. The latter can't be, since the cause can't happen after the effect. So probably updating the driver caused (or partially caused) the error messages to stop.

1. Experiments show that a person's reaction time is much longer after a few drinks but is relatively uninfluenced by a series of other factors.

2. A study showed that people with no bacteria in their mouth get no cavities – and that people with no food particles in their mouth get no cavities. However, people with both bacteria and food particles in their mouth get cavities.

3. Whenever Michelle drinks scotch and *soda*, she has a hangover the next day. Whenever she drinks gin and *soda*, she gets a hangover. Likewise, whenever she drinks rum and *soda*, she gets a hangover.

4. The morning disc jockey on a radio station remarked in early December that the coldest temperature of the day seemed to occur later and later in the morning. The weather person pointed out that the sunrise had been getting later and later. In a few weeks both processes would reverse themselves, with the sunrise and the coldest temperature of the day both occurring earlier every day.

5. Our research team at the medical center just discovered a new blood factor called "factor K." Factor K occurs in everyone who has cancer but in no one else.

6. When I sat eating on the rock slab in Grand Gulch, armies of little ants invaded the slab. Later I sat on the slab the same way except that I didn't eat anything. In the second case the ants didn't invade the slab.

7. We just did an interesting study comparing the vacation periods of employees and the disappearance of food items. We found that when Megan is working, the items disappear, and when she's away, they don't disappear.

8. People in several parts of the country have lower rates of tooth decay. Investigations show that the only thing different about these places is that their water supply contains fluoride.

9. We did an experiment where we selected two more or less identical groups and put fluoride in the first group's water but not in the second group's. The first group had a lower rate of tooth decay.

10. Many backpackers think eating raw garlic gives you an odor that causes mosquitoes not to bite you. When hiking a mosquito-infested part of the Bruce Trail, I ate much raw garlic. The mosquitoes bit me in their usual bloodthirsty manner. 0098

11. Little Will throws food on the floor and receives signs of disapproval from Mommy and Daddy. Such things happen regularly. When he eats his food without throwing it on the floor, he doesn't get any disapproval.

12. Everyone in our study who became a heroin addict had first tried marijuana.

13. If you rub two surfaces together, the surfaces get warm. They'll get warmer and warmer as you rub the surfaces together harder and faster.

14. When we plot how many hours Alex studies against the grades he gets for his various exams, we see a close correlation.

15. Matches that aren't either heated or struck don't light. Matches that are wet don't light. Matches that aren't in the presence of oxygen don't light. Matches that are heated or struck, dry, and in the presence of oxygen do light.

16. Little Will made a discovery. He keeps moving the lever on the radio up and down. He notices that the music gets louder and softer when he does this.

17. We made a careful study of the heart rate of athletes and how it correlates with various factors. The only significant correlation we found is that those who do aerobic exercise (and those alone) have lower heart rates.

18. We investigated many objects with a crystalline structure. The only thing they have in common is that all solidified from a liquid state. (Mill used this example.)

19. After long investigation, we found a close correlation between night and day. If you have night, then there invariably, in a few hours, follows day. If you have day, then invariably, in a few hours, there follows night.

20. Young Will has been experimenting with his electrical meter. He found that if he increases the electrical voltage, then he also increases the current.

21. Whenever Kurt wears his headband, he makes all his field goals. Whenever he doesn't wear it, he misses them all. This has been going on for many years.

22. The fish in my father's tank all died. We suspected either the fish food or the water temperature. We bought more fish and did everything the same except for changing the fish food. All the fish died. We then bought more fish and did everything the same except for changing the water temperature. The fish lived.

23. Bacteria introduced by visitors from the planet Krypton are causing an epidemic. We've found that everyone exposed to the bacteria gets sick and dies – except those who have a higher-than-normal heart rate.

24. When we chart the inflation rate next to the growth in the national debt over several years, we find a close correlation.

25. On my first backpack trip, I hiked long distances but wore only a single pair of socks. I got bad blisters on my feet. On my second trip, I did everything the same except that I wore two pairs of socks. I got only minor blisters.

5.8 Scientific laws

Ohm's Law is about electricity. "Law" here suggests great scientific dignity and backing. Ohm's Law is more than a mere *hypothesis* (preliminary conjecture) or even a *theory* (with more backing than a hypothesis but less than a law). 0099

Ohm's Law is a formula relating electric current, voltage, and resistance; here I = current (in amps), E = voltage (in volts), and R = resistance (in ohms):

Ohm's Law: $I = E/R$

An electric current of 1 amp (ampere) is a flow of 6,250,000,000,000,000,000 electrons per second; a 100-watt bulb draws almost an amp, and the fuse may blow if you draw over 15 amps. Voltage pushes the electrons; your outlet may have 117 volts and your flashlight battery 1.5 volts. The voltage encounters an electrical resistance, which restricts the electron flow. A short wire has low resistance (less than an ohm) while an inch of air has high resistance (billions of ohms). Small carbon resistors go from less than an ohm to millions of ohms. Ohm's Law says that current increases if you raise voltage or lower resistance.

Electric current is like the flow of water through a garden hose. Voltage is like water pressure. Electrical resistance is like the hose's resistance to the water flow; a long, thin hose has greater resistance than a short, thick one. The current or flow of water is measured in gallons per minute; it increases if you raise the water pressure or use a hose with less resistance.

Ohm's Law is a mathematical formula that lets us calculate various results. Suppose we put a 10-ohm resistor across your 117-volt electrical outlet; we'd get a current of 11.7 amps (not quite enough to blow your fuse):

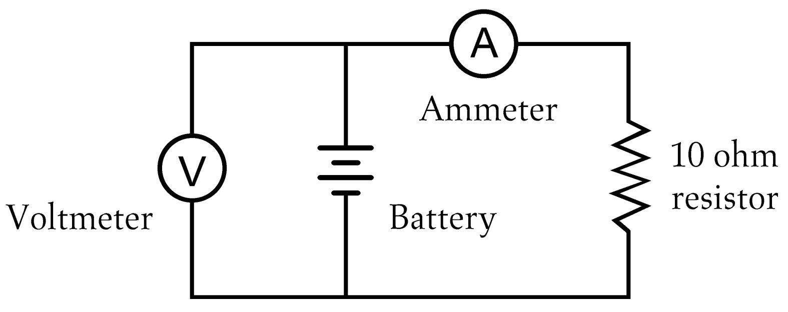
$$I = E/R = 117 \text{ volts}/10 \text{ ohms} = 11.7 \text{ amps.}$$

Ohm's Law deals with unobservable properties (current, voltage, resistance) and entities (electrons). Science allows unobservables if they have testable consequences or can somehow be measured. The term "unobservable" is vague. Actually we can feel certain voltages. The 1.5 volts from your flashlight battery can't normally be felt, slightly higher voltages give a slight tingle, and the 117 volts from your outlet can give a dangerous jolt. Philosophers dispute the status of unobservable entities. Are the ultimate elements of reality unobservables like atoms and electrons, or commonsense objects like chairs, or both? Or are atoms and chairs both just fictions to help us talk about our sensations?

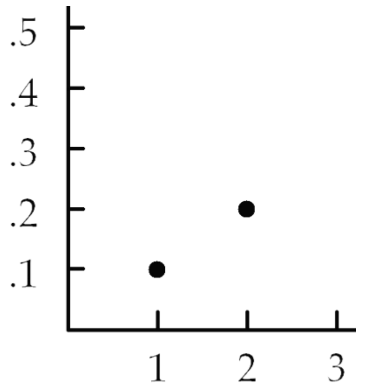
We can ask how scientific laws are *discovered*, or we can ask how they're *verified*. History can tell us how Georg Simon Ohm *discovered* his law in 1827; philosophy deals more with how such laws are *verified* (shown to be true). Roughly, scientific laws are verified by a combination of observation and argument; but the details get complicated.

Suppose we want to verify Ohm's Law. We're given batteries, resistors, and a meter for measuring current, voltage, and resistance. The meter simplifies our task; we don't have to define the fundamental units (ampere, volt, and ohm) or invent ways to measure them. Wouldn't the meter make our task too easy? Couldn't we just do a few experiments and then prove Ohm's Law, using standard deductive and inductive reasoning? Unfortunately, it's not that simple. 0100

Suppose we hook up batteries of different voltages to a resistor:



The voltmeter measures voltage, and the ammeter measures current. We start with a 10-ohm resistor. We try voltages of 1 volt and 2 volts and observe currents of .1 amp and .2 amp. Here's a chart with the results (here the horizontal x-axis is for voltage, from 0 to 3 volts, and the vertical y-axis is for current, from 0 to .5 amps):



If E = 1 volt and R = 10 ohms, then I = E/R = 1/10 = .1 amp.

If E = 2 volts and R = 10 ohms, then I = E/R = 2/10 = .2 amp.

Our observations accord with Ohm (I = E/R). So we argue inductively:

All examined voltage–resistance–current cases follow Ohm.
A large and varied group of such cases has been examined.
∴ Probably all such cases follow Ohm.

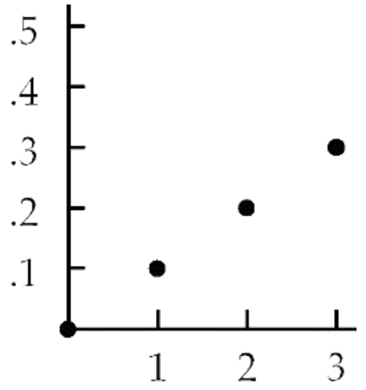
Premise 2 is weak, since we tried only two cases. But we can easily perform more experiments; after we do so, Ohm would seem to be securely based.

The problem is that we can give an inductive argument for a second and incompatible hypothesis: "I = $(E^2 - 2E + 2)/R$." Let's call this *Mho's Law*. Surprisingly, our test results also accord with Mho. In the first case (one volt), I = .1 amp [since $(1^2 - 2 \bullet 1 + 2)/10 = (1 - 2 + 2)/10 = 1/10 = .1$]; in the second case (two volts), I = .2 amp [since $(2^2 - 2 \bullet 2 + 2)/10 = (4 - 4 + 2)/10 = 2/10 = .2$]. So each examined case follows Mho. We can argue inductively as follows:

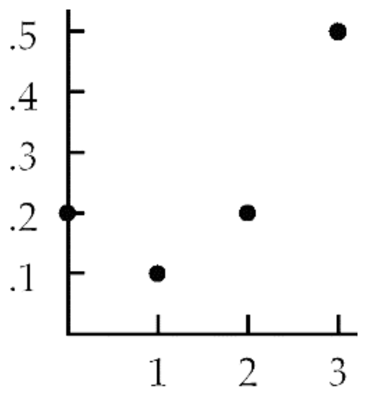
All examined voltage–resistance–current cases follow Mho.
A large and varied group of such cases has been examined.
∴ Probably all such cases follow Mho.

This inductive argument for Mho seems as strong as the one we gave for Ohm. Judging just from these arguments and test results, there seems to be no reason for preferring Ohm over Mho, or Mho over Ohm. 0101

The two laws, while agreeing on both test cases so far, give conflicting predictions for further cases. Ohm says we'll get 0 amps with 0 volts, and .3 amp with 3 volts; Mho says we'll get .2 amp with 0 volts, and .5 amp with 3 volts:



Ohm's predictions



Mho's predictions

The two laws are genuinely different, even though both give the same results for a voltage of 1 or 2 volts.

We have to try a crucial experiment to decide between the theories. What happens with 3 volts? Ohm says we'll get .3 amp, but Mho says we'll get .5 amp. If we do the experiment and get .3 amp, this would falsify Mho:

If Mho is correct and we apply 3 volts to this 10-ohm resistor, then we get .5 amp current.
We apply 3 volts to this 10-ohm resistor.
We don't get .5 amp current.
∴ Mho isn't correct.

If M and A, then G **Valid**
A
Not-G
∴ Not-M

Premise 1 links a scientific hypothesis (Mho) to antecedent conditions (that 3 volts have been applied to the 10-ohm resistor) to give a testable prediction (that we'll get .5 amp current). Premise 2 says the antecedent conditions have been fulfilled. But premise 3 says the results conflict with what was predicted. Since this argument has true premises and is deductively valid, our experiment shows Mho to be wrong.

Does our experiment similarly show that Ohm is correct? Unfortunately not. Consider this argument:

If Ohm is correct and we apply 3 volts to this 10-ohm resistor, then we get .3 amp current.
We apply 3 volts to this 10-ohm resistor.
We get .3 amp current.
∴ Ohm is correct.

If O and A, then G **Invalid**
A
G
∴ O

This is invalid, as we could check using propositional logic (Chapter 6). So the premises don't prove that Ohm is correct; and Ohm might fail for further cases. But the experiment strengthens our inductive argument for Ohm, since it gives a larger and more varied sample. So we can have greater trust that the pattern observed to hold so far will continue to hold.

Here are three aspects of scientific method: 0102

- Scientists often set up crucial experiments to decide between conflicting theories. Scientists dream up alternative theories and look for ways to decide between them.
- *We can sometimes deductively refute a theory through a crucial experiment.* Experimental results, when combined with other suitable premises, can logically entail that a theory is false.
- *We can't deductively prove a theory using experiments.* Experiments can

> inductively support a theory and deductively refute opposing theories. But they can't eliminate the possibility of the theory's failing for further cases.

Recall how the Mho problem arose. We had two test cases that agreed with Ohm. These test cases also agreed with another formula, one we called "Mho"; and the inductive argument for Mho seemed as strong as the one for Ohm. But Ohm and Mho gave conflicting predictions for further test cases. So we did a crucial experiment to decide between the two. Ohm won.

There's always another Mho behind the bush – so our problems aren't over. However many experiments we do, there are always alternative theories that agree with all test cases so far but disagree on some further predictions. In fact, there's always an *infinity* of theories that do this. No matter how many dots we put on the chart (representing test results), we could draw an unlimited number of lines that go through all these dots but otherwise diverge.

Suppose we conduct 1000 experiments in which Ohm works. There are alternative theories Pho, Qho, Rho, and so on that agree on these 1000 test cases but give conflicting predictions about further cases. And each theory seems to be equally supported by the same kind of inductive argument:

All examined voltage–resistance–current cases follow this theory.
A large and varied group of such cases has been examined.
∴ Probably all such cases follow this theory.

Even after 1000 experiments, Ohm is just one of infinitely many formulas that seem equally probable on the basis of the test results and inductive logic.

In practice, we prefer Ohm on the basis of *simplicity*. Ohm is the simplest formula that agrees with all our test results. So we prefer Ohm to the alternatives and see Ohm as firmly based.

What is simplicity and how can we decide which of two scientific theories is simpler? We don't have neat and tidy answers to these questions. In practice, though, we can tell that Ohm is simpler than Mho; we judge that Ohm's formula and straight line are simpler than Mho's formula and curved line. We don't have a clear and satisfying definition of "simplicity"; yet we can apply this notion in a rough way in many cases.

The simplicity criterion is a form of *Ockham's razor* (§16.2): 0103

> **Simplicity criterion**: Other things being equal, we ought to prefer a simpler theory to a more complex one.

The "other things being equal" qualification is important. Experiments may force us to accept very complex theories; but we shouldn't accept such theories unless we have to.

It's unclear how to justify the simplicity criterion. Since inductive reasoning stumbles unless we presuppose the criterion, an inductive justification would be circular. Perhaps the criterion is a self-evident truth not in need of justification. Or perhaps it's pragmatically justified:

> If the simplicity criterion isn't correct, then no scientific laws are justified.
> Some scientific laws are justified.
> ∴ The simplicity criterion is correct.

Does premise 2 beg the question against the skeptic? Can this premise be defended without appealing to the criterion? The simplicity criterion is vague and raises complex problems, but we can't do without it.

Coherence is another factor that's important for choosing between theories:

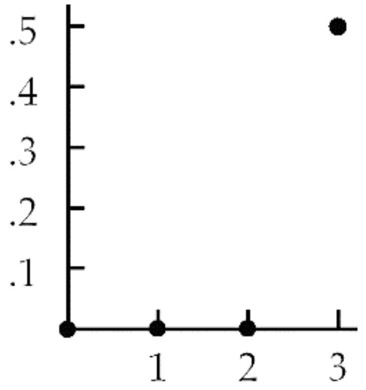
> **Coherence criterion**: Other things being equal, we ought to prefer a theory that harmonizes with existing well-established beliefs.

Mho has trouble here, since it predicts that 0 volts across a 10-ohm resistor produces a .2 amp current. But then it follows, using an existing well-grounded belief that current through a resistor produces heat, that a 10-ohm resistor with no voltage applied produces heat. While nice for portable handwarmers, this would be difficult to harmonize with the conservation of energy. So the coherence criterion leads us to doubt Mho.

Do further tests continue to confirm Ohm? The answer is complicated. Some resistors give, not a straight-line chart, but a curve; this happens if we use an incandescent light bulb for the resistor. Instead of rejecting Ohm, scientists say that heating the resistor changes the resistance. This seems satisfactory, since the curve becomes straighter if the resistor is kept cooler. And we can measure changes in resistance when the resistor is heated externally.

Another problem is that resistors will burn up or explode if enough voltage is applied. This brings an irregularity into the straight-line chart. But again, scientists regard this as changing the resistance, and not as falsifying Ohm.

A more serious problem is that some devices don't even roughly match the pattern predicted by Ohm. A Zener diode, for example, draws almost no current until a critical voltage is reached; then it draws a high current: 0104



Zener diode curve

Do such devices refute Ohm? Not necessarily. Scientists implicitly qualify Ohm so it applies just to "pure resistances" and not to things like Zener diodes. This seems circular. Suppose that a "pure resistor" is any device that satisfies Ohm. Then isn't it circular to say that Ohm holds for "pure resistors"? Doesn't this just mean that Ohm works for any device for which it works?

In practice, people working in electronics quickly learn which devices satisfy Ohm and which don't. The little tubular "resistors" follow Ohm closely (neglecting slight changes caused by heating and major changes when we burn up the resistor). Zener diodes, transistors, and other semiconductors generally don't follow Ohm. So Ohm can be a useful principle, even though it's difficult to specify in any precise and non-circular manner the cases where it applies.

5.8a Exercise

Sketch in a rough way how we might verify or falsify these hypotheses. Point out any special difficulties likely to arise.

Women have less innate logical ability than men.
We'd give a logic test to large and varied groups of either sex, and see how results differ. If women tested lower [they don't – judging from a test I designed for a friend in psychology], this wouldn't itself prove lower innate ability, since the lower scores might come from different social expectations or upbringing. It would be difficult to avoid this problem completely; but we might try testing groups in cultures with less difference in social expectations and upbringing.

1. Neglecting air resistance, objects of any weight fall at the same speed.
2. Germs cause colds.

3. A huge Ice-Age glacier covered most of Wisconsin about 10,000 years ago.

4. Regular moderate use of marijuana is no more harmful than regular moderate use of alcohol.

5. When couples have several children, the child born first tends to have greater innate intelligence than the one born last.

6. Career-oriented women tend to have marriages that are more successful than those of home-oriented women.

7. Factor K causes cancer. 0105

8. Water is made up of molecules consisting of two atoms of hydrogen and one atom of oxygen.

9. Organisms of a given biological species randomly develop slightly different traits; organisms with survival-promoting traits tend to survive and pass these traits to their offspring. New biological species result when this process continues over millions of years. This is how complex species developed from simple organisms, and how humans developed from lower species.

10. Earth was created 5,000 years ago, complete with all current biological species.

5.9 Best-explanation reasoning

Suppose you made fudge for a party. When you later open the refrigerator, you find that most of the fudge is gone. You also find that your young son, who often steals deserts, has fudge on his face. The child denies that he ate the fudge. He contends that Martians appeared, ate the fudge, and spread some on his face. But you aren't fooled. The better and more likely explanation is the child ate the fudge. So this is what you believe.

This is an **inference to the best explanation**. Intuitively, we should accept the best explanation for the data. Consider what we said about Ohm's Law in the previous section. Ohm's Law explains a wide range of phenomena about electrical voltage, current, and resistance. Besides having testable implications that accord well with our experience, the law also has other virtues, including clarity, simplicity, and coherence with existing well-established beliefs. Unless someone comes up with a better explanation of the data, we should accept Ohm's Law.

Our best argument for the theory of evolution has a similar form:

We ought to accept the best explanation for the wide range of empirical facts about biological organisms (including comparative structure, embryology, geographical distribution, and fossil records).
The best explanation for the wide range of empirical facts about biological organisms is evolution.
∴ We ought to accept evolution.

A fuller formulation would elaborate on what these empirical facts are, alternative ways to explain them, and why evolution provides a better explanation than its rivals. Some think our core beliefs about most things, including the existence of material objects, other minds, and perhaps God, are to be justified as inferences to the best explanation.

Particularly interesting is the "fine-tuning" inference for the existence of God. Here the empirical data to be explained is that the basic physical constants that govern the universe (like the gravitational constant "g," the charge and mass of the proton, the density of water, and the total mass of the universe) are within the very narrow range that makes it possible for life to evolve. Stephen Hawking 0106 gives this example: "If the rate of expansion one second after the Big Bang had been smaller by even one part in a hundred thousand million million, the universe would have recollapsed before it ever reached its present size"[1] – which would have blocked the evolution of life. So life requires the expansion rate to be correct to the 17th decimal place; and other constants are similar. How is this empirical data to be explained? Could this precise combination of physical constants have come about by chance? Some atheists propose that there are an infinity of parallel universes, each governed by a different physics, and that it was highly likely that *some* of these parallel universes could produce life. But many theists claim that the simplest and best explanation involves God: that the universe was caused by a great mind who "fine tuned" its physical laws to make possible the emergence of life.

The general form of the inference to the best explanation raises some issues. On what grounds should we evaluate one explanation as "better" than another? Should we accept the *best possible* explanation (even though no one may yet have thought of it) or the *best currently available* explanation (even though none of the current explanations may be very good)? And why is the best explanation most likely to be the true one?

[1] *A Brief History of Time*, tenth anniversary edition (New York: Bantam Books, 1998), page 126; he also gives other examples and discusses their theological implications. Anthony Flew (*There Is a God* (New York: HarperCollins, 2007), pp. 113–21) and Francis S. Collins (*The Language of God* (New York: Free Press, 2006), pp. 63–84) were prominent atheists who converted to theism due to the fine-tuning argument. Http://www.harryhiker.com/reason.pdf defends the argument and http://www.harryhiker.com/genesis.exe is a corresponding Windows computer game.

5.10 Problems with induction

We've seen that inductive logic isn't as neat and tidy as deductive logic. Now we'll consider two further perplexing problems: how to *formulate* principles of inductive logic and how to *justify* these principles.

We've formulated inductive principles in rough ways that if taken literally can lead to absurdities. For example, our statistical-syllogism formulation can lead to this absurd inference:

60 percent of all Chicago voters are Democrats.
This non-Democrat is a Chicago voter.
That's all we know about the matter.
∴ It's 60 percent probable that this non-Democrat is a Democrat.

Actually, "This non-Democrat is a Democrat" is 0 percent probable, since it's a self-contradiction. So our statistical syllogism principle isn't entirely correct.

We noted that the analogy syllogism is oversimplified in its formulation. We need to rely on *relevant similarities* instead of just counting resemblances. But 0107 "relevant similarities" is hard to pin down.

Sample-projection syllogisms suffer from a problem raised by Nelson Goodman. Consider this argument:

All examined diamonds are hard.
A large and varied group of diamonds has been examined.
∴ Probably all diamonds are hard.

Given that the premises are true, the argument would seem to be a good one. But consider this second argument, which has the same form except that we substitute a more complex phrase for "hard":

All examined diamonds are such that they are hard-if-and-only-if-they-were-examined-before-the-year-2222.
A large and varied group of diamonds has been examined.
∴ Probably all diamonds are such that they are hard-if-and-only-if-they-were-examined-before-the-year-2222.

Premise 1 is tricky to understand. It's not yet 2222. So if all examined diamonds are hard, then they are such that they are hard-if-and-only-if-they-were- examined-before-the-year-2222. So premise 1 is true. Premise 2 also is true. Then this second argument also would seem to be a good one.

Consider a diamond X that will first be examined after 2222. By our first argument, diamond X probably *is* hard; by the second, it probably *isn't* hard. So our sample projection argument leads to conflicting conclusions.

Philosophers have discussed this problem for decades. Some suggest that we qualify the sample-projection syllogism form to outlaw the second argument; but it's unclear how to eliminate the bad apples without also eliminating the good ones. As yet, there's no agreement on how to solve the problem.

Goodman's problem is somewhat like one we saw in the last section. Here we had similar inductive arguments for two incompatible laws: Ohm and Mho:

All examined electrical cases follow Ohm's Law.
A large and varied group of cases has been examined.
∴ Probably all electrical cases follow Ohm's Law.

All examined electrical cases follow Mho's Law.
A large and varied group of cases has been examined.
∴ Probably all electrical cases follow Mho's Law.

Even after 1000 experiments, there still are an *infinity* of theories that give the same test results in these 1000 cases but conflicting results in further cases. And we could "prove," using an inductive argument, that each of these incompatible theories is probably true. But this is absurd. We can't have each of an infinity of conflicting theories be probably true. Our sample-projection syllogism thus leads to absurdities. 0108

We got around this problem in the scientific-theory case by appealing to simplicity: "Other things being equal, we ought to prefer a *simpler* theory to a more complex one." While "simpler" here is vague and difficult to explain, we seem to need some such simplicity criterion to justify any scientific theory.

Simplicity is important in our diamond case, since 1 is simpler than 2:

1. All diamonds are hard.
2. All diamonds are such that they are hard-if-and-only-if-they-were-examined-before-the-year-2222.

By our simplicity criterion, we ought to prefer 1 to 2, even if both have equally strong inductive backing. So the sample-projection syllogism seems to need a simplicity qualification too; but it's not clear how to formulate it.

So it's difficult to formulate clear inductive-logic principles that don't lead to absurdities. Inductive logic is less neat and tidy than deductive logic.

Our second problem is how to justify inductive principles. For now, let's ignore the problem we just talked about. Let's pretend that we have clear inductive principles that roughly accord with our practice and don't lead to absurdities. Why follow these principles?

Consider this inductive argument (which says roughly that the sun will probably come up tomorrow, since it has come up every day in the past):

> All examined days are days in which the sun comes up.
> A large and varied group of days has been examined.
> Tomorrow is a day.
> ∴ Probably tomorrow is a day in which the sun comes up.

Even though the sun has come up every day in the past, it still might not come up tomorrow. Why think that the premise gives good reason for accepting the conclusion? Why accept this or any inductive argument?

David Hume several centuries ago raised this problem about the justification of induction. We'll discuss five responses.

1. Some suggest that, to justify induction, we need to presume that nature is uniform. If nature works in regular patterns, then the cases we haven't examined will likely follow the same patterns as the ones we have examined.

There are two problems with this suggestion. First, what does it mean to say "Nature is uniform"? Let's be concrete. What would this principle imply about the regularity (or lack thereof) of Chicago weather patterns? "Nature is uniform" seems either very vague or clearly false.

Second, what's the backing for the principle? Justifying "Nature is uniform" by experience would require inductive reasoning. But then we're arguing in a circle – using the uniformity idea to justify induction, and then using induction to justify the uniformity idea. This presumes what's being doubted: that it's reasonable to follow inductive reasoning in the first place. Or is the uniformity idea perhaps a self-evident truth not in need of justification? But it's 0109 implausible to claim self-evidence for a claim about what the world is like.

2. Some suggest that we justify induction by its success. Inductive methods work. Using inductive reasoning, we know what to do for a toothache and how to fix cars. We use such reasoning continuously and successfully in our lives. What better justification for inductive reasoning could we have than this?

This seems like a powerful justification. But there's a problem. Let's assume that inductive reasoning has worked in the past; how can we then conclude that it probably will work in the future? The argument is inductive, much like our sunrise argument:

Induction has worked in the past.
∴ Induction probably will work in the future.

The sun has come up every day in the past.
∴ The sun probably will come up tomorrow.

So justifying inductive reasoning by its past success is circular; it uses inductive reasoning and thus presupposes that such reasoning is legitimate.

3. Some suggest that it's part of the meaning of "reasonable" that beliefs based on inductive reasoning are *reasonable*. "Reasonable belief" just means "belief based on experience and inductive reasoning." So it's true by definition that beliefs based on experience and inductive reasoning are reasonable.

There are two problems with this. First, the definition is wrong. It really isn't true by definition that all and only things based on experience and inductive reasoning are reasonable. There's no contradiction in disagreeing with this – as there would be if this definition were correct. Mystics see their higher methods as reasonable, and skeptics see the ordinary methods as unreasonable. Both groups might be wrong, but they aren't simply contradicting themselves.

Second, even the correctness of the definition wouldn't solve the problem. Suppose that standards of inductive reasoning are built into the conventional meaning of our word "reasonable." Suppose that "reasonable belief" simply means "belief based on experience and inductive reasoning." Then why follow what's "reasonable" in this sense? Why not instead follow the skeptic's advice and avoid believing such things? So this semantic approach doesn't answer the main question: Why follow inductive reasoning at all?

4. Karl Popper suggests that we avoid inductive reasoning. But we seem to need such reasoning in our lives; without inductive reasoning, we have no basis for believing that bread nourishes and arsenic kills. And suggested substitutes for inductive reasoning don't seem adequate.

5. Some suggest that we approach justification in inductive logic the same way we approach it in deductive logic. How can we justify the validity of deductive principles like *modus ponens* ("If A then B, A ∴ B")? Can we prove such principles? Perhaps we can prove *modus ponens* by doing a truth table (§6.6) and then arguing this way: 0110

If the truth table for *modus ponens* never gives true premises and a false conclusion, then *modus ponens* is valid.
The truth table for *modus ponens* never gives true premises and a false conclusion.
∴ *Modus ponens* is valid.

Premise 1 is a necessary truth and premise 2 is easy to check. The conclusion follows. Therefore, *modus ponens* is valid. But the problem is that the argument itself uses *modus ponens*. So this attempted justification is circular, since it presumes from the start that *modus ponens* is valid.

Aristotle long ago showed that every proof must eventually rest on something unproved; otherwise, we'd need an infinite chain of proofs or else circular arguments – and neither is acceptable. So why not just accept the validity of *modus ponens* as a self-evident truth – a truth that's evident but can't be based on anything more evident? If we have to accept some things as evident without proof, why not accept *modus ponens* as evident without proof?

I have some sympathy with this approach. But, if we accept it, we shouldn't think that picking logical principles is purely a matter of following "logical intuitions." Logical intuitions vary enormously among people. The pretest that I give shows that most beginning logic students have poor intuition about the validity of simple arguments. But even though untrained logical intuitions differ, still we can reach agreement on many principles of logic. Early on, we introduce the notion of logical form. And we distinguish between valid and invalid forms – such as these two:

Modus ponens

If A then B **Valid**
A
∴ B

Affirming the consequent

If A then B **Invalid**
B
∴ A

Students at first are poor at distinguishing valid from invalid forms. They need concrete examples like these:

If you're a dog, then you're an animal. **Valid**
You're a dog.
∴ You're an animal.

If you're a dog, then you're an animal. **Invalid**
You're an animal.
∴ You're a dog.

After enough well-chosen examples, the validity of *modus ponens* and the invalidity of affirming the consequent become clear.

So, despite the initial clash of intuitions, we eventually reach clear logical principles of universal rational appeal. We do this by searching for clear formulas that lead to intuitively correct results in concrete cases without leading to any clear absurdities. We might think that this procedure proves *modus ponens*: 0111

> If *modus ponens* leads to intuitively correct results in concrete cases without leading to any clear absurdities, then *modus ponens* is valid.
> *Modus ponens* leads to intuitively correct results in concrete cases without leading to any clear absurdities.
> ∴ *Modus ponens* is valid.

But this reasoning itself uses *modus ponens*; the justification is circular, since it presumes from the start that *modus ponens* is valid. So this procedure of testing *modus ponens* by checking its implications doesn't prove *modus ponens*. But I think it gives a "justification" for it, in some sense of "justification." This is vague, but I don't know how to make it more precise.

I suggested that we justify inductive principles the same way we justify deductive ones. Realizing that we can't prove everything, we wouldn't demand a proof. Rather, we'd search for clear formal inductive principles that lead to intuitively correct results in concrete cases without leading to any clear absurdities. Once we reached such inductive principles, we'd rest content with them and not look for any further justification.

This is the approach that I'd use in justifying inductive principles. But the key problem is the one discussed earlier. As yet we seem unable to find clear formal inductive principles that lead to intuitively correct results in concrete cases without leading to any clear absurdities. We just don't know how to formulate inductive principles very rigorously. This is what makes the current state of inductive logic intellectually unsatisfying.

Inductive reasoning has been very useful. Inductively, we assume that it will continue to be useful. In our lives, we can't do without it. But the intellectual basis for inductive reasoning is shaky.

6 Basic Propositional Logic

Propositional logic studies arguments whose validity depends on "if-then," "and," "or," "not," and similar notions. This chapter covers the basics and the next covers proofs. Our later logical systems build on what we learn here.

6.1 Easier translations

We'll now create a "propositional language," with precise rules for constructing arguments and testing validity. Our language uses capital letters for true-or-false statements, parentheses for grouping, and five special *logical connectives* ("∼" squiggle, "•" dot, "∨" vee, "⊃" horseshoe, and "≡" threebar):

∼P	=	Not-P
(P • Q)	=	Both P and Q
(P ∨ Q)	=	Either P or Q
(P ⊃ Q)	=	If P then Q
(P ≡ Q)	=	P if and only if Q

A grammatically correct formula of our language is called a **wff**, or **well-formed formula**. Wffs are sequences that we can construct using these rules:[1]

1. Any capital letter is a wff.
2. The result of prefixing any wff with "∼" is a wff.
3. The result of joining any two wffs by "•" or "∨" or "⊃" or "≡" and enclosing the result in parentheses is a wff.

These rules let us build wffs like the following:

P
= I live in Paris.

∼Q

[1] Pronounce "wff" as "woof" (as in "wood"). We'll take letters with primes (like **A′** and **A′′**) to be additional letters.

= I don’t live in Quebec.

$(P \bullet {\sim}Q)$
= I live in Paris and I don’t live in Quebec.

$(N \supset (P \bullet {\sim}Q))$
= If I’m Napoleon, then I live in Paris and not Quebec. 0113

“$\sim$P” doesn’t need or use parentheses. A wff requires a pair of parentheses for each “$\bullet$,” “$\vee$,” “$\supset$,” or “$\equiv$.” So “$\sim P \bullet Q$” is malformed and not a wff; this ambiguous formula could be given parentheses in two ways:

$({\sim}P \bullet Q)$ = Both not-P and Q
${\sim}(P \bullet Q)$ = Not both P and Q

The first says definitely that P is false and Q is true. The second just says that not both are true (at least one is false). Don’t read both the same way, as “not P and Q.” Read “both” for the left-hand parenthesis, or use pauses:

$({\sim}P \bullet Q)$ = Not-P (pause) and (pause) Q
${\sim}(P \bullet Q)$ = Not (pause) P and Q

Logic is easier if you read the formulas correctly. These two also differ:

$(P \bullet (Q \supset R))$ = P, and if Q then R
$((P \bullet Q) \supset R)$ = If P-and-Q, then R

The first says P is definitely true, but the second leaves us in doubt about this.

Here’s a useful rule for translating from English into logic, with examples:

Put “(” wherever you see “both,” “either,” or “if.”

Either not A or B = $({\sim}A \vee B)$
Not either A or B = ${\sim}(A \vee B)$
If both A and B, then C = $((A \bullet B) \supset C)$
Not both not A and B = ${\sim}({\sim}A \bullet B)$

Our translation rules have exceptions and need to be applied with common sense. So don’t translate “I saw them both” as “S(” – which isn’t a wff.

Here’s another rule:

Group together parts on either side of a comma.

If A, then B and C $=$ $(A \supset (B \bullet C))$
If A then B, and C $=$ $((A \supset B) \bullet C)$

If you’re confused on where to divide a sentence without a comma, ask yourself where a comma would naturally go, and then translate accordingly:

If it snows then I’ll go outside and I’ll ski
$=$ $(S \supset (G \bullet K))$
If it snows, then I’ll go outside and I’ll ski

Be sure that your capital letters stand for whole statements. “Gensler is happy” is just “G”; don’t use “$(G \bullet H)$” (“Gensler *and* happy”?). Similarly, “Bob and Lauren got married to each other” is just “M”; “$(B \bullet L)$” would be wrong, since the English sentence doesn’t mean “Bob got married and Lauren got married” (which omits “to each other”). However, it would be correct to translate 0114 “Bob and Lauren were sick” as “$(B \bullet L)$”; here “and” connects whole statements since the English means “Bob was sick and Lauren was sick.”

It doesn’t matter what letters you use, as long as you’re consistent. Use the same letter for the same idea and different letters for different ideas. If you use “P” for “I went to Paris,” then use “$\sim P$” for “I didn’t go to Paris.”

Order and grouping don’t matter in wffs using “$\bullet$,” “$\vee$,” or “$\equiv$” as the only connective:[1]

$(A \bullet B)$ $=$ $(B \bullet A)$
$((A \bullet B) \bullet C)$ $=$ $(A \bullet (B \bullet C))$

Order matters with “$\supset$”; these two make different claims:

If it’s a dog, then it’s an animal $=$ $(D \supset A)$
If it’s an animal, then it’s a dog $=$ $(A \supset D)$

We can switch the parts of an if-then if we negate them; so “If it’s a dog, then it’s an animal” “$(D \supset A)$” is equivalent to the *contrapositive* “If it’s not an animal, then it’s not a dog” “$(\sim A \supset \sim D)$.”

[1] Order matters in English when “and” means “and then”; “Suzy got married and had a baby” differs from “Suzy had a baby and got married.” Our “$\bullet$” is simpler and more abstract, and ignores temporal sequence. §§7.5 and 15.2 have additional equivalences.

6.1a Exercise: LogiCola C (EM & ET)[1]

Translate these English sentences into wffs.

Both not A and B.
(∼A • B)

1. Not both A and B.
2. Both A and either B or C.
3. Either both A and B or C.
4. If A, then B or C.
5. If A then B, or C.
6. If not A, then not either B or C.
7. If not A, then either not B or C.
8. Either A or B, and C.
9. Either A, or B and C.
10. If A then not both not B and not C.
11. If you get an error message, then the disk is bad or it's a Macintosh disk.
12. If I bring my digital camera, then if my batteries don't die then I'll take pictures of my backpack trip and put the pictures on my Web site.
13. If you both don't exercise and eat too much, then you'll gain weight. 0115
14. The statue isn't by either Cellini or Michelangelo.
15. If I don't have either $2 in exact change or a bus pass, I won't ride the bus.
16. If Michigan and Ohio State play each other, then Michigan will win.
17. Either you went through both Dayton and Cinci, or you went through Louisville.
18. If she had hamburgers then she ate junk food, and she ate French fries.
19. I'm going to Rome or Florence and you're going to London.
20. Everyone is male or female.

[1] Exercise sections have a boxed sample problem that's worked out. They also refer to LogiCola computer exercises (see Preface), which give a fun and effective way to master the material. Problems 1, 3, 5, 10, 15, and so on are worked out in the answer section at the back of the book.

6.2 Basic truth tables

Let "P" stand for "I went to Paris" and "Q" for "I went to Quebec." Each could be *true* or *false* (the two **truth values**) – represented by "1" and "0" (or sometimes "T" and "F"). There are four possible combinations:

P Q	
0 0	Both are false
0 1	Just Q is true
1 0	Just P is true
1 1	Both are true

I went to neither Paris nor Quebec
I went to Quebec but not Paris
I went to Paris but not Quebec
I went to both Paris and Quebec

A **truth table** gives a logical diagram for a wff. It lists all possible truth-value combinations for the letters and says whether the wff is true or false in each case. The truth table for "•" ("and") is very simple:

P Q	(P • Q)
0 0	0
0 1	0
1 0	0
1 1	1

"I went to Paris and I went to Quebec."

"(P • Q)" is a **conjunction**; P and Q are its **conjuncts**.

"(P • Q)" claims that *both* parts are true. So "I went to Paris *and* I went to Quebec" is false in the first three cases (where one or both parts are false) – and true only in the last case. These truth equivalences give the same information:

(0 • 0) = 0 (false • false) = false
(0 • 1) = 0 (false • true) = false
(1 • 0) = 0 (true • false) = false
(1 • 1) = 1 (true • true) = true

"$(0 \bullet 0) = 0$" says that an AND statement is false if both parts are false. The next two say that an AND is false if one part is false and the other part is true. And "$(1 \bullet 1) = 1$" says that an AND is true if both parts are true.

Here are the truth table and equivalences for "∨" ("or"): 0116

P Q	(P ∨ Q)
0 0	0
0 1	1
1 0	1
1 1	1

$$(0 \vee 0) = 0$$
$$(0 \vee 1) = 1$$
$$(1 \vee 0) = 1$$
$$(1 \vee 1) = 1$$

"I went to Paris or I went to Quebec."

"(P ∨ Q)" is a **disjunction**; P and Q are its **disjuncts.**

"(P ∨ Q)" claims that *at least one* part is true. So "I went to Paris *or* I went to Quebec" is true just if I went to one or both places. Our "∨" symbolizes the *inclusive* sense of "or"; English also can use "or" in an *exclusive* sense, which claims that at least one part is true *but not both*:

Inclusive "or": A or B or both = (A ∨ B)
Exclusive "or": A or B but not both = ((A ∨ B) • ~(A • B))

So the exclusive sense requires a longer symbolization.[1]

Here are the truth table and equivalences for "⊃" ("if-then"):

P Q	(P ⊃ Q)
0 0	1
0 1	1
1 0	0
1 1	1

$$(0 \supset 0) = 1$$
$$(0 \supset 1) = 1$$
$$(1 \supset 0) = 0$$

[1] People sometimes use "*Either* A *or* B" for the exclusive "or." We won't do this; instead, we'll use "either" to indicate grouping and we'll translate it as a left-hand parenthesis.

$(1 \supset 1) = 1$

"If I went to Paris, then I went to Quebec."

"$(P \supset Q)$" is a **conditional**; P is the **antecedent** and Q the **consequent.**

"$(P \supset Q)$" claims that what we *don't* have is the first part true and the second false. Suppose you say this:

"*If* I went to Paris, *then* I went to Quebec."

By our table, you speak truly if you went to neither place, or to both places, or to Quebec but not Paris. You speak falsely if you went to Paris but not Quebec. Does that seem right to you? Most people think so, but some have doubts.

Our truth table can produce strange results. Take this example:

If I had eggs for breakfast, then the world will end at noon = $(E \supset W)$

Suppose I didn't have eggs, and so E is false. By our table, the conditional is then *true* – since if E is false then "$(E \supset W)$" is true. This is strange. We'd normally take the conditional to be *false* – since we'd take it to claim that my having eggs would *cause* the world to end. So translating "if-then" as "⊃" seem fishy.

Our "⊃" symbolizes a simplified "if-then" that ignores causal connections and temporal sequence. "$(P \supset Q)$" has a very simple meaning; it just *denies* that we have P-true-and-Q-false: 0117

$(P \supset Q)$ = $\sim(P \bullet \sim Q)$

If P is true, then Q is true = We don't have P true and Q false

Translating "if-then" this way is a useful simplification, since it captures the part of "if-then" that normally determines validity. The simplification usually works; in the few cases where it doesn't, we can use a more complex translation (as we'll sometimes do in the chapters on modal logic).

The truth conditions for "⊃" are hard to remember. These slogans may help:

Falsity implies anything.
$(0 \supset \quad) = 1$

Anything implies truth.
$(\quad \supset 1) = 1$

Truth doesn't imply falsity.
$(1 \supset 0) = 0$

"Falsity implies anything," for example, means that the whole if-then is true if the first part is false; so "If I'm a billionaire, then …" is true, regardless of what replaces "…," since I'm *not* a billionaire.

Here are the table and equivalences for "≡" ("if-and-only-if"):

P Q	(P ≡ Q)
0 0	1
0 1	0
1 0	0
1 1	1

$(0 \equiv 0) = 1$
$(0 \equiv 1) = 0$
$(1 \equiv 0) = 0$
$(1 \equiv 1) = 1$

"I went to Paris if and only if I went to Quebec."

"(P ≡ Q)" is a **biconditional.**

"(P ≡ Q)" claims that both parts have the *same* truth value: both are true or both are false. So "≡" is much like "equals."

Here are the table and equivalences for "∼" ("not"):

P	∼P
0	1
1	0

$\sim 0 = 1$
$\sim 1 = 0$

"I didn't go to Paris."

"∼P" is a **negation.**

"∼P" has the *opposite* value of "P." If "P" is true then "∼P" is false, and if "P" is false then "~P" is true.

This double-box sums up these **basic truth equivalences** (learn them well!):

AND	OR	NOT
$(0 \bullet 0) = 0$ $(0 \bullet 1) = 0$ $(1 \bullet 0) = 0$ $(1 \bullet 1) = 1$	$(0 \vee 0) = 0$ $(0 \vee 1) = 1$ $(1 \vee 0) = 1$ $(1 \vee 1) = 1$	$\sim 0 = 1$ $\sim 1 = 0$
both parts are true	*at least one* part is true	*reverse* the truth value

IF-THEN	IFF
$(0 \supset 0) = 1$ $(0 \supset 1) = 1$ $(1 \supset 0) = 0$ $(1 \supset 1) = 1$	$(0 \equiv 0) = 1$ $(0 \equiv 1) = 0$ $(1 \equiv 0) = 0$ $(1 \equiv 1) = 1$
we don't have *first true & second false*	both parts have *same* truth value

0118

6.2a Exercise: LogiCola D (TE & FE)

Calculate each truth value.

$(0 \bullet 1)$
$(0 \bullet 1) = 0$

1. $(0 \vee 1)$
2. $(0 \bullet 0)$
3. $(0 \supset 0)$
4. ~ 0
5. $(0 \equiv 1)$
6. $(1 \bullet 0)$
7. $(1 \supset 1)$
8. $(1 \equiv 1)$
9. $(0 \vee 0)$
10. $(0 \supset 1)$
11. $(0 \equiv 0)$

12. (1 ∨ 1)
13. (1 • 1)
14. (1 ⊃ 0)
15. ∼1
16. (1 ∨ 0)
17. (1 ≡ 0)

6.3 Truth evaluations

We can calculate a wff's truth value if we know the truth value of its letters:

Suppose that P = 1, Q = 0, and R = 0. What's the truth value of "((P ⊃ Q) ≡ ∼R)"?

First replace "P" with "1" and the other letters with "0," to get "((1 ⊃ 0) ≡ ∼0)." Then simplify from the inside out, using our basic truth equivalences, until we get "1" or "0." Here we get "0," so the formula is false:

Formula: ((1 ⊃ 0) ≡ ∼0)
Replace "(1 ⊃ 0)" with "0" and "∼0" with "1," to get "(0 ≡ 1)"
Replace "(0 ≡ 1)" with "0," to get "0"

In evaluating "((1 ⊃ 0) ≡ ∼0)," we keep looking for parts, here highlighted as "((1 ⊃ 0) ≡ ∼0)," that match the left side of our basic truth equivalences (see previous page), and then replace these parts with their equivalents.

On this strategy, with formulas like "∼(1 ∨ 0)," first work out the truth value of the part in parentheses. Then apply "∼" to the result:

Formula: ∼(1 ∨ 0)
Replace "(1 ∨ 0)" with "1," to get "∼1"
Replace "∼1" with "0," to get "0"

Beginners often do this wrong. They distribute the NOT, going from "∼(1 ∨ 0)" to "(∼1 ∨ ∼0)" (wrong!); this evaluates to "(0 ∨ 1)" and then "1" (wrong!). *Don't distribute "NOT"!* With "∼(...)," first simplify the part in parentheses and then apply "∼" to the result.[1] 0119

[1] NOT ("∼") doesn't distribute in logic, since, for example, "∼(P • Q)" (which says that not both are true) differs from "(∼P • ∼Q)" (which says that both are false). Likewise MINUS ("–") doesn't distribute in math, since "–(2 • 2)" (which equals –4) differs from "(–2 • –2)" (which equals +4).

6.3a Exercise: LogiCola D (TM & TH)

Assume that A = 1 and B = 1 (A and B are both true) while X = 0 and Y = 0 (X and Y are both false). Calculate the truth value of each wff below.

((A ∨ X) ⊃ ∼B)
((1 ∨ 0) ⊃ ∼1) (1 ⊃ 0) 0

1. ∼(A • X)
2. (∼A • ∼X)
3. ∼(∼A • ∼X)
4. (A ⊃ X)
5. (∼X ≡ Y)
6. (∼B ⊃ A)
7. ∼(A ⊃ X)
8. (B • (X ∨ A))
9. (∼(X • A) ∨ ∼B)
10. (∼A ∨ ∼(X ⊃ Y))
11. ((A • ∼X) ⊃ ∼B)
12. ∼(A ⊃ (X ∨ ∼B))
13. (∼X ∨ ∼(∼A ≡ B))
14. (∼Y ⊃ (A • X))
15. ∼((A ⊃ B) ⊃ (B ⊃ Y))

6.4 Unknown evaluations

We can often figure out a formula's truth value without knowing the value of some letters:

> Suppose that P = 1 and Q = ? (unknown). What's the truth value of "(P ∨ Q)"?

We might just see that "(1 ∨ ?)" is true, since an OR is true if at least one part is true. Or we can try it both ways; "(1 ∨ **?**)" is *true* because it's true either way:

$$(1 \lor \mathbf{1}) \;=\; \mathbf{1}$$
$$(1 \lor \mathbf{0}) \;=\; \mathbf{1}$$

Here's another example:

> Suppose that P = 1 and Q = ? What is the truth value of "(P • Q)"?

We might just see that "(1 • ?)" is unknown, since its truth value depends on the unknown letter. Or we can try it both ways; "(1 • **?**)" is *unknown* because it could turn out true and it could turn out false:

$$(1 \bullet \mathbf{1}) \;=\; \mathbf{1}$$
$$(1 \bullet \mathbf{0}) \;=\; \mathbf{0}$$

6.4a Exercise: LogiCola D (UE, UM, & UH)

Assume that T = 1 (T is true), F = 0 (F is false), and U = ? (U is unknown). Calculate the truth value of each wff below. 0120

(∼T • U)
(∼1 • ?) = (0 • ?) = 0

1. (U • F)
2. (U ⊃ ∼T)
3. (U ∨ ∼F)
4. (∼F • U)
5. (F ⊃ U)
6. (∼T ∨ U)
7. (U ⊃ ∼T)
8. (∼F ∨ U)
9. (T • U)
10. (U ⊃ ∼F)
11. (U • ∼T)
12. (U ∨ F)

6.5 Complex truth tables

A truth table for a wff is a diagram listing all possible truth-value combinations for the wff's letters and saying whether the wff would be true or false in each case. We've done simple tables already; now we'll do complex ones.

With n distinct letters we have 2^n possible truth-value combinations. And so one letter gives 2 (2^1) combinations:

A
0
1

Two letters give 4 (2^2) combinations:

A	B
0	0
0	1
1	0
1	1

Three letters give 8 (2^3) combinations:

A	B	C
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

And n letters give 2^n combinations. To get every combination, alternate 0's and 1's for the last letter the required number of times. Then alternate 0's and 1's for each earlier letter at half the previous rate: by twos, fours, and so on. This numbers the rows in base 2.

Begin a truth table for "∼(A ∨ ∼B)" like this:

A B	∼(A ∨ ∼B)

0 0	
0 1	
1 0	
1 1	

The right side has the wff. The left side has each letter used in the wff; write each letter just once, regardless of how often it occurs. Below the letters, write all possible truth-value combinations. Then figure out the wff's truth value for each line. The first line has A and B both false – which makes the whole wff false:

Formula: ∼(A ∨ ∼B)
Replace each letter with "0," to get "∼(0 ∨ ∼0)"
Replace "∼0" with "1," to get "∼(0 ∨ 1)"
Replace "(0 ∨ 1)" with "1," to get "∼1"
Replace "∼1" with "0," to get "0" 0121

The wff comes out "1," "0," and "0" for the next three lines; so we get:

A B	∼(A ∨ ∼B)
0 0	0
0 1	1
1 0	0
1 1	0

"∼(A ∨ ∼B)" is true if and only if A is false and B is true. The simpler wff "(∼A • B)" is equivalent, in that it's true in the same cases. Both wffs are true in some cases and false in others – making them *contingent statements*.

"(P ∨ ∼P)" is a *tautology*, since it comes out true in all cases:

P	(P ∨ ∼P)
0	1
1	1

"I went to Paris or I didn't go to Paris."

This formula, the **law of the excluded middle**, says that every statement is true or false. This holds in propositional logic, since we stipulated that capital letters stand for true-or-false statements. The law doesn't always hold in English, since English allows statements that are too vague to be true or false, like "It's raining" when there's a slight drizzle or "My shirt is white" when it's a light cream color. So the law is an idealization when applied to English.

"(P • ~P)" is a **self-contradiction**, since it comes out false in all cases:

P	(P • ~P)
0	0
1	0

"I went to Paris and I didn't go to Paris."

"P and not-P" is always false in propositional logic, which presupposes that "P" stands for the same statement throughout. English is looser and lets us shift the meaning of a phrase in the middle of a sentence. "I went to Paris and I didn't go to Paris" may express a truth if it means "I went to Paris (in that I landed once at the Paris airport) – but I didn't really go there (in that I saw almost nothing of the city)." Because of the shift in meaning, this better translates as "(P • ~Q)."

6.5a Exercise: LogiCola D (FM & FH)

Do a truth table for each wff. 0122

((P ∨ Q) ⊃ R)	
P Q R	((P ∨ Q) ⊃ R)
0 0 0	1
0 0 1	1
0 1 0	0
0 1 1	1
1 0 0	0
1 0 1	1
1 1 0	0
1 1 1	1

1. (P ≡ ~Q)
2. (~P • Q)
3. (P ∨ (Q • ~R))
4. ((P • ~Q) ⊃ R)
5. ((P ≡ Q) ⊃ Q)
6. ((P ∨ ~Q) ⊃ R)
7. (~Q ⊃ ~P)

8. (P ≡ (P • P))

9. ~(P • (Q ∨ ~R))

6.6 The truth-table test

Recall how we defined VALID and INVALID for arguments:

VALID = No possible case has premises all true and conclusion false. *This can't happen: 1, 1 ∴ 0*	
INVALID = Some possible case has premises all true and conclusion false. *This can happen: 1, 1 ∴ 0*	

To use the *truth-table test* on a propositional argument:

> Construct a truth table showing the truth value of the premises and conclusion for all possible cases. The argument is **valid** if and only if no possible case has premises all true and conclusion false.

Suppose we want to test this invalid argument; first do a truth table for premises and conclusion, starting as follows:

If you're a dog, then you're an animal.
You're not a dog.
∴ You're not an animal.

(D ⊃ A)
~D
∴ ~A

D	A	(D ⊃ A),	~D	∴	~A
0	0				
0	1				
1	0				
1	1				

Then evaluate the three wffs on each truth combination. The first combination 0123 has D = 0 and A = 0, which makes all three wffs true:

$(D \supset A) = (0 \supset 0) = 1$
$\sim D = \sim 0 = 1$
$\sim A = \sim 0 = 1$

So the first line of our truth table looks like this:

D A	(D ⊃ A),	~D	∴	~A
0 0	1	1		1

Work out the other three lines:

D A	(D ⊃ A),	~D	∴	~A
0 0	1	1		1
0 1	**1**	**1**		**0**
1 0	0	0		1
1 1	1	0		0

Invalid – we can get true premises and a false conclusion (second line).

The argument is invalid, since some possible case has premises all true and conclusion false. Perhaps you're an animal but not a dog (but maybe a cat).

With this next argument, again do a truth table for premises and conclusion:

If you're a dog, then you're an animal.
You're a dog.
∴ You're an animal.

(D ⊃ A)
D
∴ A

D A	(D ⊃ A),	D	∴	A
0 0	1	0		0
0 1	1	0		1
1 0	0	1		0
1 1	1	1		1

Valid – we never get true premises and a false conclusion.

There’s a short-cut test. Recall that we’re looking for 110 (premises all true and conclusion false). The argument is invalid if 110 sometimes occurs; otherwise, it’s valid. To save time, first evaluate an easy wff and cross out lines that can’t be 110. In our last example, we might work out “D” first:

D A	(D ⊃ A),	D	∴	A
0 0	-----	0		---
0 1	-----	0		---
1 0		1		
1 1		1		

The first two lines can’t be 110 (since the second digit is 0); so we cross them out and ignore them. Next we might evaluate “A”: 0124

D A	(D ⊃ A),	D	∴	A
0 0	-----	0		---
0 1	-----	0		---
1 0		1		0
1 1	-----	1		1

The bottom line can’t be 110 (since the last digit is 1); so we cross it out. Then we evaluate “(D ⊃ A)” for only one case – for which it comes out false. Since we never get 110, the argument is valid:

D A	(D ⊃ A),	D	∴	A
0 0	-----	0		---
0 1	-----	0		---
1 0	0	1		0
1 1	-----	1		1

Valid – we never get true premises and a false conclusion.

The short-cut method can save much time if otherwise we’d have to evaluate a long formula for eight or more cases.

With a two-premise argument, look for 110. With three premises, look for 1110. In general, look for a case having premises all true and conclusion false. The argument is valid if and only if this never occurs.

The truth-table test can get tedious for long arguments. Arguments with 6 letters need 64 lines – and ones with 10 letters need 1024 lines. So we'll use the truth-table test only on fairly simple arguments.[1]

6.6a Exercise: LogiCola D (AE, AM, & AH)

First appraise intuitively. Then translate into logic (using the letters given) and use the truth-table test to determine validity.

It's in my left hand or my right hand.
It's not in my left hand.
∴ It's in my right hand.

L R	(L ∨ R),	~L	∴	R
0 0	0	1		0
0 1	1	1		1
1 0	1	0		0
1 1	1	0		1

Valid – we never get true premises & false conclusion.

1. If you're a collie, then you're a dog.
You're a dog.
∴ You're a collie. [Use C and D.]

2. If you're a collie, then you're a dog.
You're not a dog.
∴ You're not a collie. [Use C and D.] 0125

3. If television is always right, then Anacin is better than Bayer.
If television is always right, then Anacin isn't better than Bayer.
∴ Television isn't always right. [Use T and B.]

4. If it rains and your tent leaks, then your down sleeping bag will get wet.
Your tent won't leak.
∴ Your down sleeping bag won't get wet. [R, L, W]

[1] An argument that tests out "invalid" may be valid on grounds that go beyond the system in question. For example, "This is green, therefore something is green" translates into propositional logic as "T ∴ S" and tests out invalid; but it's valid as "Gt ∴ (∃x)Gx" in quantificational logic.

5. If I get Grand Canyon reservations and get a group together, then I'll explore canyons during spring break.
I've got a group together.
I can't get Grand Canyon reservations.
∴ I won't explore canyons during spring break. [R, T, E]

6. There's an objective moral law.
If there's an objective moral law, then there's a source of the moral law.
If there's a source of the moral law, then there's a God. (Other possible sources, like society or the individual, are claimed not to work.)
∴ There's a God. [Use M, S, and G; from C. S. Lewis.]

7. If ethics depends on God's will, then something is good because God desires it.
Something isn't good *because* God desires it. (Instead, God desires something because it's already good.)
∴ Ethics doesn't depend on God's will. [Use D and B; from Plato's *Euthyphro*.]

8. It's an empirical fact that the basic physical constants are precisely in the narrow range of what is required for life to be possible. (This "fine-tuning principle" has considerable evidence behind it.)
The best explanation for this fact is that the basic physical constants were caused by a great mind intending to produce life. (The main alternatives are the "chance coincidence" and "parallel universe" explanations.)
If these two things are true, then it's reasonable to believe that the basic structure of the world was set up by a great mind (God) intending to produce life.
∴ It's reasonable to believe that the basic structure of the world was set up by a great mind (God) intending to produce life. [Use E, B, and R; see §5.9.]

9. I'll go to Paris during spring break if and only if I'll win the lottery.
I won't win the lottery.
∴ I won't go to Paris during spring break. [P, W]

10. If we have a simple concept proper to God, then we've directly experienced God and we can't rationally doubt God's existence.
We haven't directly experienced God.
∴ We can rationally doubt God's existence. [S, E, R]

11. If there is a God, then God created the universe.
If God created the universe, then matter didn't always exist.
Matter always existed.
∴ There is no God. [G, C, M] 0126

12. If this creek is flowing, then either the spring upstream has water or this creek has some other water source.
This creek has no other water source.
This creek isn't flowing.
∴ The spring upstream has no water.   [F, S, O]

6.7 The truth-assignment test

Recall how we defined VALID and INVALID for arguments:

VALID = No possible case has premises all true and conclusion false. *This can't happen: 1, 1 ∴ 0*	
INVALID = Some possible case has premises all true and conclusion false. *This can happen: 1, 1 ∴ 0*	

To use the *truth-assignment test* on a propositional argument:

> Set each premise to 1 and the conclusion to 0. Figure out the truth value of as many letters as possible. The argument is **valid** if and only if no possible way to assign 1 and 0 to the letters will keep the premises all 1 and conclusion 0.

Suppose we want to test this valid argument:

It's in my left hand or my right hand.
It's not in my left hand.
∴ It's in my right hand.

(L ∨ R)
~L
∴ R

Here's how we work it out. First set each premise to 1 and the conclusion to 0:

$(L \vee R) = 1$
$\sim L = 1$
$\therefore R = 0$

Since premise 2 has $\sim L = 1$, making $L = 0$, write 0 above each L. A 0 superscript above a letter, as in "$\sim L^0$," says that that letter is false:

$(L^0 \vee R) = 1$
$\sim L^0 = 1$
$\therefore R = 0$

Since the conclusion has $R = 0$, write 0 above each R: 0127

$(L^0 \vee R^0) = 1$
$\sim L^0 = 1$
$\therefore R^0 = 0$

But then premise 1 can't be true. So we can't have true premises and a false conclusion. So it's **valid**:

$(L^0 \vee R^0) \neq 1$ **Valid**
$\sim L^0 = 1$
$\therefore R^0 = 0$

So first assign 1 to the premises and 0 to the conclusion (just to see if this could work). Then figure out the truth values for the letters, and then for the longer formulas. If we have to cross something out, then the initial assignment isn't possible, and so the argument is valid.

This next example shows how to work out an invalid argument:

It's in my left hand or my right hand.
It's not in my left hand.
∴ It's not in my right hand.

$(L \vee R)$
$\sim L$
$\therefore \sim R$

First set each premise to 1 and the conclusion to 0:

$(L \vee R) = 1$
$\sim L = 1$

$\therefore {\sim}R \ = \ 0$

Since premise 2 has ~L = 1, making L = 0, write 0 above each L:

$(L^0 \vee R) \ = \ 1$
${\sim}L^0 \ = \ 1$
$\therefore {\sim}R \ = \ 0$

Since the conclusion has ~R = 0, making R = 1, write 1 above each R:

$(L^0 \vee R^1) \ = \ 1$
${\sim}L^0 \ = \ 1$
$\therefore {\sim}R^1 \ = \ 0$

So we can have true premises and a false conclusion. So it's **invalid**:

$(L^0 \vee R^1) \ = \ 1$ **Invalid**
${\sim}L^0 \ = \ 1$
$\therefore {\sim}R^1 \ = \ 0$

A truth table gives the same result when L = 0 and R = 1:

L R	(L ∨ R),	~L	∴	~R
0 1	**1**	**1**		**0**

Invalid

The truth-assignment test gives this result more quickly.
Here's another invalid argument:

It's in my left hand or my right hand.
∴ It's in my right hand.

$(L \vee R)$
$\therefore R$

If we work this out, we get R false, but we get no value for L; so we give L a value that makes all premises true and conclusion false. Again, first set the premise to 1 and the conclusion to 0: 0128

$(L \vee R) \ = \ 1$
$\therefore R \ = \ 0$

Since the conclusion has R=0, write 0 above each R:

$(L \vee R^0) \ = \ 1$
$\therefore R^0 \ = \ 0$

To make the premise true, make L true:

$(L^1 \vee R^0) \ = \ 1$
$\therefore R^0 \ = \ 0$

So we can have true premises and a false conclusion. So it's INVALID:

$(L^1 \vee R^0) \ = \ 1$ **Invalid**
$\therefore R^0 \ = \ 0$

If you don't get a value for a letter, try it both ways (as 1 and as 0); if either gives true premises and a false conclusion, then the argument is invalid.

In working out the truth values for the letters, try to make premises all true and conclusion false. The argument is invalid if there's some way to do this.

6.7a Exercise: LogiCola ES

Test for validity using the truth-assignment test.

$(K \supset (I \vee S))$ $\sim I$ K $\therefore S$
$(K^1 \supset (I^0 \vee S^0)) \ \neq \ 1$ **Valid** $\sim I^0 \ = \ 1$ $K^1 \ = \ 1$ $\therefore S^0 \ = \ 0$ (we can't have 1110)

1. $\sim(N \equiv H)$
N
$\therefore \sim H$

2. ((J • ~D) ⊃ Z)
~Z
D
∴ ~J

3. ((T ∨ M) ⊃ Q)
M
∴ Q

4. P
∴ (P • Q)

5. ((L • F) ⊃ S)
S
F
∴ L

6. ((A • U) ⊃ ~B)
B
A
∴ ~U

7. ((W • C) ⊃ Z)
~Z
∴ ~C

8. Q
∴ (P ⊃ Q)

9. (E ∨ (Y • X))
~E
∴ X

10. (~T ⊃ (P ⊃ J))
P
~J
∴ T

11. ~P
∴ ~(Q ⊃ P)

12. ((~M • G) ⊃ R)
~R
G
∴ M

13. ~(Q ≡ I)
~Q
∴ I

14. ((Q • R) ≡ S)
Q
∴ S

15. A
~A
∴ B 0129

6.7b Exercise: LogiCola EE

First appraise intuitively. Then translate into logic and use the truth-assignment test to determine validity.

If our country will be weak, then there will be war.
Our country will not be weak.
∴ There will not be war.

$(K^0 \supset R^1) = 1$ **Invalid**
$\sim K^0 = 1$
$\therefore \sim R^1 = 0$

(we can have 110)

1. Some things are caused (brought into existence).
Anything caused is caused by another.
If some things are caused and anything caused is caused by another, then either there's a first cause or there's an infinite series of past causes.
There's no infinite series of past causes.
∴ There's a first cause. [A "first cause" (often identified with God) is a cause that isn't itself caused by another; from St Thomas Aquinas.]

2. If you pass and it's intercepted, then the other side gets the ball.
You pass.
It's not intercepted.
∴ The other side doesn't get the ball.

3. If God exists in the understanding and not in reality, then there can be conceived a being greater than God (namely, a similar being that also exists in reality).
"There can be conceived a being greater than God" is false (since "God" is defined as "a being than which no greater can be conceived").
God exists in the understanding.
∴ God exists in reality. [This is St Anselm's famous ontological argument.]

4. If existence is a perfection and God by definition has all perfections, then God by definition must exist.
Existence is a perfection.
God by definition has all perfections.
∴ God by definition must exist. [From René Descartes.]

5. If we have sensations of alleged material objects and yet no material objects exist, then God is a deceiver.
God isn't a deceiver.
We have sensations of alleged material objects.
∴ Material objects exist. [From René Descartes, who thus based our knowledge of the external material world on our knowledge of God.]

6. If "good" is definable in experimental terms, then ethical judgments are scientifically provable and ethics has a rational basis.
Ethical judgments aren't scientifically provable.
∴ Ethics doesn't have a rational basis. 0130

7. If it's right for me to lie and not right for you, then there's a relevant difference between our cases.
There's no relevant difference between our cases.
It's not right for you to lie.
∴ It's not right for me to lie.

8. If Newton's gravitational theory is correct and there's no undiscovered planet near Uranus, then the orbit of Uranus would be such-and-such.
Newton's gravitational theory is correct.
The orbit of Uranus isn't such-and-such.
∴ There's an undiscovered planet near Uranus. [This reasoning led to the discovery of the planet Neptune.]

9. If attempts to prove "God exists" fail in the same way as our best arguments for "There are other conscious beings besides myself," then belief in God is reason- - able if and only if belief in other conscious beings is reasonable.
Attempts to prove "God exists" fail in the same way as our best arguments for "There are other conscious beings besides myself."
Belief in other conscious beings is reasonable.
∴ Belief in God is reasonable. [From Alvin Plantinga.]

10. If you pack intelligently, then either this teddy bear will be useful on the hiking trip or you won't pack it.
This teddy bear won't be useful on the hiking trip.
You won't pack it.
∴ You pack intelligently.

11. If knowledge is sensation, then pigs have knowledge.
Pigs don't have knowledge.
∴ Knowledge isn't sensation. [From Plato.]

12. If capital punishment is justified and justice doesn't demand a vindication for past wrongs, then capital punishment reforms the offender or effectively deters crime.
Capital punishment doesn't reform the offender.
Capital punishment doesn't effectively deter crime.
∴ Capital punishment isn't justified.

13. If belief in God were a purely intellectual matter, then either all smart people would be believers or all smart people would be non-believers.
Not all smart people are believers.
Not all smart people are non-believers.
∴ Belief in God isn't a purely intellectual matter.

14. If you're lost, then you should call for help or head downstream.
You're lost.
∴ You should call for help. 0131

15. If maximizing human enjoyment is always good and the sadist's dog-torturing maximizes human enjoyment, then the sadist's act is good.
The sadist's dog-torturing maximizes human enjoyment.
The sadist's act isn't good.
∴ Maximizing human enjoyment isn't always good.

16. If there's knowledge, then either some things are known without proof or we can prove every premise by previous arguments infinitely.
We can't prove every premise by previous arguments infinitely.
There's knowledge.
∴ Some things are known without proof. [From Aristotle.]

17. If you modified your computer or didn't send in the registration card, then the warranty is void.
You didn't modify your computer.
You sent in the registration card.
∴ The warranty isn't void.

18. If "X is good" means "Hurrah for X!" and it makes sense to say "If X is good," then it makes sense to say "If hurrah for X!"
It makes sense to say "If X is good."
It doesn't make sense to say "If hurrah for X!"
∴ "X is good" doesn't mean "Hurrah for X!" [From Hector-Neri Castañeda.]

19. If we have an idea of substance, then "substance" refers either to a simple sensation or to a complex constructed out of simple sensations.
"Substance" doesn't refer to a simple sensation.
∴ We don't have an idea of substance. [From David Hume.]

20. If we have an idea of "substance" and we don't derive the idea of "substance" from sensations, then "substance" is a thought category of pure reason.
We don't derive the idea of "substance" from sensations.
We have an idea of "substance."
∴ "Substance" is a thought category of pure reason. [From Immanuel Kant.]

21. If "good" means "socially approved," then what is socially approved is necessarily good.
What is socially approved isn't necessarily good.
∴ "Good" doesn't mean "socially approved."

22. [Generalizing the last argument, G. E. Moore argued that we can't define "good" in terms of any empirical term "F" – like "desired" or "socially approved."]
If "good" means "F," then what is F is necessarily good.
What is F isn't necessarily good. (We can consistently say "Some F things may not be good" without thereby violating the meaning of "good.")
∴ "Good" doesn't mean "F."

23. If moral realism (the belief in objective moral truths) were true, then it could explain the moral diversity in the world.
Moral realism can't explain the moral diversity in the world.
∴ Moral realism isn't true. 0132

6.8 Harder translations

As you symbolize idiomatic English, keep following our earlier rules: (1) put "(" wherever you see "both," "either," or "if"; and (2) group together parts on either side of a comma. Here we'll add additional rules, with examples:

Translate "but" ("yet," "however," "although," and so on) as "and."

Michigan played *but* it lost
= $(P \bullet L)$

The translation loses the contrast (or surprise), but this doesn't affect validity.

Translate "unless" as "or."

You'll die *unless* you breathe
= $(D \vee B)$ = $(B \vee D)$

Unless you breathe you'll die
= $(D \vee B)$ = $(B \vee D)$

"Unless" is also equivalent to "if not"; so we also could use "$(\sim B \supset D)$" ("If you don't breathe, then you'll die").

Translate "just if" and "iff" (a logician word) as "if and only if."

I'll agree *just if* you pay me $1,000
= $(A \equiv P)$

I'll agree *iff* you pay me $1,000
= $(A \equiv P)$

The order of the letters doesn't matter with "•" or "∨" or "≡."

Our next two rules are tricky. The first governs most conditional words:

The part after "if" ("provided that," "assuming that," and so on) is the if-part (the antecedent, the part before the horseshoe).

If A, then B
= (A ⊃ B)

Provided that A, B
= (A ⊃ B)

A, if B
= (B ⊃ A)

A, provided that B
= (B ⊃ A)

You're an animal, *if* you're a dog
= (D ⊃ A)

Provided that you're a dog, you're an animal
= (D ⊃ A)

"Only if" is different and follows its own rule:

The part after "only if" is the then-part (the consequent, the part after the horseshoe). (Or just write "⊃" for "only if.")

A only if B
= (A ⊃ B)

Only if A, B
= (B ⊃ A) 0133

You're alive *only if* you have oxygen
= (A ⊃ O)

Only if you have oxygen, are you alive
= (A ⊃ O)

The *contrapositive* translation "(∼O ⊃ ∼A)" ("If you don't have oxygen, then you aren't alive") is equivalent and often sounds more intuitive.

Here's the rule for "sufficient" and "necessary":

"A is *sufficient* for B" means "If A then B." "A is *necessary* for B" means "If not A then not B." "A is *necessary and sufficient* for B" means "A if and only if B."

Water is *sufficient for* life
= $(W \supset L)$

Water is *necessary for* life
= $(\sim W \supset \sim L)$

Water is necessary and sufficient for life
= $(W \equiv L)$

The order of the letters matters with "⊃," but not with "≡."

Sometimes none of these rules applies and you just have to puzzle out the meaning on your own.

6.8a Exercise: LogiCola C (HM & HT)

Translate these English sentences into wffs.

A, assuming that B.
$(B \supset A)$

1. If she goes, then you'll be alone but I'll be here.
2. Your car will start only if you have fuel.
3. I will quit unless you give me a raise.
4. Taking the final is a sufficient condition for passing.
5. Taking the final is necessary for you to pass.
6. You're a man just if you're a rational animal.
7. Unless you have faith, you'll die.
8. She neither asserted it nor hinted at it.
9. Getting at least 96 is a necessary and sufficient condition for getting an A.

10. Only if you exercise are you fully alive.

11. I'll go, assuming that you go.

12. Assuming that your belief is false, you don't know.

13. Having a true belief is a necessary condition for having knowledge.

14. You get mashed potatoes or French fries, but not both.

15. You're wrong if you say that. 0134

6.9 Idiomatic arguments

Our arguments so far have been phrased in a clear premise–conclusion format. Unfortunately, real-life arguments are seldom so neat and clean. Instead we often find convoluted wording or extraneous material. Important parts of the argument may be omitted or only hinted at. And it may be hard to pick out the premises and conclusion. It often takes hard work to reconstruct a clearly stated argument from a passage.

Logicians like to put the conclusion (here italicized) last:

"Socrates is human. If he's human, then he's mortal. *So Socrates is mortal*

H
(H ⊃ M)
∴ M

But people sometimes put the conclusion first, or in the middle:

"*Socrates must be mortal.* After all, he's human. And if he's human, he's mortal."

"Socrates is human. *So he must be mortal* – since if he's human, he's mortal."

Here "must" and "so" indicate the conclusion (which always goes last when we translate into logic). Here are some typical words that help us pick out premises and conclusion:

These often indicate premises:

Because, for, since, after all …
I assume that, as we know …
For these reasons …

These often indicate conclusions:

Hence, thus, so, therefore …
It must be, it can't be …
This proves (or shows) that …

When you don't have this help, ask yourself what is argued *from* (these are the premises) and what is argued *to* (this is the conclusion).

In reconstructing an argument, first pick out the conclusion. Then symbolize the premises and conclusion; this may involve untangling idioms like "A unless B" (which translates as "A or B"). If you don't get a valid argument, try adding unstated but implicit premises (you may need to add a premise that uses letters that only occur once); using the "principle of charity," interpret unclear reasoning in the way that gives the best argument.

Here's a twisted argument – and how it goes into premises and a conclusion:

The gun must have been shot recently! It's still hot.

The gun is still hot.
∴ The gun was shot recently.

H
∴ S 0135

Since this seems to presume an implicit premise, we add a plausible one that makes the argument valid. Then we translate into logic and test for validity:

If the gun is still hot, then it was shot recently. (implicit)
The gun is still hot.
∴ The gun was shot recently.

(H ⊃ S) **Valid**
H
∴ S

6.9a Exercise: LogiCola E (F & I)

First appraise intuitively. Then pick out the conclusion, translate into logic, and determine validity using the truth-assignment test. Supply implicit premises if needed.

Knowledge is good in itself only if it's desired for its own sake. So knowledge is good in itself, since it's desired for its own sake.
$(G^0 \supset D^1) = 1$ **Invalid** $D^1 = 1$ $\therefore G^0 = 0$ The conclusion is "*So knowledge is good in itself*": "G."

1. Knowledge can't be sensation. If it were, then we couldn't know something that we aren't presently sensing. [From Plato.]

2. Presuming that we followed the map, then unless the map is wrong there's a pair of lakes just over the pass. We followed the map. There's no pair of lakes just over the pass. Hence the map is wrong.

3. If they blitz but don't get to our quarterback, then our wide receiver will be open. So our wide receiver won't be open, as shown by the fact that they won't blitz.

4. My true love will marry me only if I buy her a Rolls–Royce. It follows that she'll marry me, since I'll buy her a Rolls–Royce.

5. The basic principles of ethics can't be self-evident truths, since if they were then they'd largely be agreed upon by intelligent people who have studied ethics.

6. That your views are logically consistent is a necessary condition for your views to be sensible. Your views are logically consistent. So your views are sensible.

7. If Ohio State wins but Nebraska doesn't, then the Ohio Buckeyes will be national champions. So it looks like the Ohio Buckeyes won't be national champs, since Nebraska clearly is going to win.

8. The filter capacitor can't be blown. This is indicated by the following facts. You'd hear a hum, presuming that the silicon diodes work but the filter capacitor is blown. But you don't hear a hum. And the silicon diodes work.

9. There's oxygen present. And so there will be a fire! My reason for saying this is that only if there's oxygen present will there be a fire.

10. We have no moral knowledge. This is proved by the fact that if we did have moral knowledge then basic moral principles would be either provable or self-evident. But they aren't provable. And they aren't self-evident either.

11. It must be a touchdown! We know that it's a touchdown if the ball broke the plane of the end zone. 0136

12. Assuming that it wasn't an inside job, then the lock was forced unless the thief stole the key. The thief didn't steal the key. We may infer that the robbery was an inside job, inasmuch as the lock wasn't forced.

13. It must be the case that we don't have any tea bags. After all, we'd have tea bags if your sister Carol drinks tea. Of course, Carol doesn't drink tea.

14. We can't still be on the right trail. We'd see the white Appalachian Trail blazes on the trees if we were still on the right trail.

15. If God is omnipotent, then he could make hatred inherently good – unless there's a contradiction in hatred being inherently good. But there's no contradiction in this. And God is omnipotent. I conclude that God could make hatred inherently good. [From William of Ockham, who saw morality as depending on God's will.]

16. Taking the exam is a sufficient condition for getting an A. You didn't take the exam. This means you don't get an A.

17. If Texas or Arkansas wins, then I win my $10 bet. I guess I win $10. Texas just beat Oklahoma 17–14!

18. Unless you give me a raise, I'll quit. Therefore I'm quitting!

19. Empirical knowledge must be impossible. My reason for saying this is that there's no independent way to prove that our senses are reliable. Empirical knowledge would be possible, of course, only if there were an independent way to prove that our senses are reliable.

20. It's virtuous to try to do what's good. On the other hand, it's not virtuous to try to do what's socially approved. I conclude that, contrary to cultural relativism, "good" doesn't mean "socially approved." I assume, of course, that if "good" meant "socially approved" and it was virtuous to try to do what's good, then it would be virtuous to try to do what's socially approved.

21. Moral conclusions can be deduced from non-moral premises only if "good" is definable using non-moral predicates. But "good" isn't so definable. So moral conclusions can't be deduced from non-moral premises.

22. The world can't need a cause. If the world needed a cause, then so would God.

6.10 S-rules

Inference rules are rules of valid reasoning that provide the building blocks for formal proofs (which we begin in the next chapter). We'll name our inference rules after the type of wff that they operate on, like AND or IF-THEN.

S-rules *simplify* statements. Our first S-rule simplifies AND statements and is itself called "AND":

AND	(P • Q) –––––– P, Q
AND statement, so both parts are true.	

From an AND statement, we can infer each part: “It’s cold and windy; therefore it’s cold, therefore it’s windy.” Negative parts work the same way: 0137

It’s not cold and it’s not windy.
∴ It’s not cold.
∴ It’s not windy.

(∼C • ∼W)
–––––––––
∼C, ∼W

But from a NOT-BOTH statement (where “∼” is outside the parentheses), we can infer nothing about the truth or falsity of the parts:

You’re *not both* in Paris and in Quebec.
∴ No conclusion.

∼(P • Q)
–––––––
nil

From “∼(P • Q)” we can’t tell the truth value for “P” or for “Q”; we only know that not both are true (at least one is false). Use the AND rule only on AND forms, like these three:

(A • B)
(∼C • D)
((E ≡ F) • (G ∨ H))

Never use the AND rule on a formula that starts with a NOT, like “∼(J • K)”; this formula has the NOT-BOTH form, not the AND form. ANDs always start with “(” and then have a wff and “•” and a wff and “)”; ANDs never start with a squiggle (“∼”).

Our second S-rule operates on NOR (NOT-EITHER) statements and is itself called “NOR”:

<table>
<tr><td>NOR</td><td>∼(P ∨ Q)
––––––––
∼P, ∼Q</td></tr>
<tr><td colspan="2">NOT-EITHER is true, so both parts are false.</td></tr>
</table>

From a NOR, we can infer the opposite of each part: “It’s not either cold or windy, therefore it’s not cold, therefore it’s not windy.” Negative parts work the same way: infer the opposite of each part (the opposite of “∼A” being “A”):

Not either not-A or not-B.
∴ A
∴ B

∼(∼A ∨ ∼B)
———————
A, B

∼(part-1 ∨ part-2)
—————————————
op of part-1, op of part-2

But a *positive* OR tells us nothing about the truth or falsity of each part:

You’re in either Paris or Quebec.
∴ No conclusion.

(P ∨ Q)
————
nil

Here we can’t tell the truth or falsity of each part; we only know that at least one part is true. Use the NOR rule only on NOR forms, like these three below:

∼(A ∨ B)
∼(∼C ∨ D)
∼((E ≡ F) ∨ (G • H))

0138 NORs always start with a squiggle. Never use the NOR rule on an OR, like “(J ∨ K).”

Our final S-rule operates on NIF (FALSE IF-THEN) statements:

NIF	∼(P ⊃ Q) ————— P, ∼Q
FALSE IF-THEN, so first part true, second part false.	

Since “(P ⊃ Q)” means “We *don’t* have P-true-and-Q-false,” so also “∼(P ⊃ Q)” means “We *do* have P-true-and-Q-false.” NIF isn’t very intuitive; memorize it instead of appealing to intuitions or examples. You’ll use this rule so much in doing proofs that it’ll become second nature.

If a NIF has negative parts, again infer part-1 and the opposite of part-2:

$$\frac{\sim(\sim A \supset B)}{\sim A, \sim B}$$

$$\frac{\sim(A \supset \sim B)}{A, B}$$

$$\frac{\sim(\sim A \supset \sim B)}{\sim A, B}$$

$$\frac{\sim(\text{part-1} \supset \text{part-2})}{\text{part-1, op of part-2}}$$

A positive IF-THEN “$(A \supset B)$” says nothing about each part’s truth or falsity. Use the NIF rule only on NIF (FALSE IF-THEN) forms, like these:

$$\sim(A \supset B)$$
$$\sim(\sim C \supset D)$$
$$\sim((E \equiv F) \supset (G \bullet H))$$

NIFs always start with a squiggle. Never use the NIF rule on an IF-THEN, like “$(J \supset K)$.”

And so you can simplify AND, NOR, and NIF:

AND	NOR	NIF
$\frac{(P \bullet Q)}{P, Q}$	$\frac{\sim(P \vee Q)}{\sim P, \sim Q}$	$\frac{\sim(P \supset Q)}{P, \sim Q}$
“AND statement, so both parts are true.”	“NOT-EITHER is true, so both parts are false.”	“FALSE IF-THEN, so first part true, second part false.”

I suggest that, as you apply these rules, you mumble the little saying at the bottom – like “AND statement, so both parts are true.” To understand why our rules work, recall our basic truth tables:

- A true AND must have both parts true.

- A false OR must have both parts false.
- A false IF-THEN must have the first part true and the second part false.

Try to learn the inference rules so well that they become automatic. You'll use these rules a lot when you do formal proofs; and learning formal proofs will be so much easier if you've already mastered the inference rules. 0139

6.10a Exercise: LogiCola F (SE & SH)

Draw any simple conclusions (a letter or its negation) that follow from these premises. If nothing follows, leave blank.

(C • ~R) --------
(C • ~R) -------- C, ~R "AND statement, so both parts are true."

1. (P • U)
   ----

2. (L ∨ C)
   ----

3. (~N ⊃ S)
   ----

4. ~(F ⊃ M)
   ----

5. ~(R ∨ S)
   ----

6. ~(J • ~N)
   ----

7. ~(I ∨ ~V)
   ----

8. (F ⊃ ~G)
   ----

9. (~Q • B)
   ----

10. ~(H ⊃ ~I)

11. (~O ∨ ~X)

12. (~T ⊃ ~H)

13. ~(~N ∨ ~E)

14. ~(Q • T)

15. (M ∨ ~W)

16. (~D • ~Z)

17. ~(~Y ⊃ G)

18. ~(~A • ~J)

19. ~(~U ⊃ ~L)

20. (~K ∨ B)

6.11 I-rules

I-rules *infer* a conclusion from two premises. Our first I-rule is called "NOT-BOTH," since the larger wff has to have this form:

NOT-BOTH

~(P • Q) P —————— ~Q	~(P • Q) Q —————— ~P

affirm one part

NOT-BOTH are true, this one is, so the other one isn't.

To infer with NOT-BOTH, we must *affirm* one part:

You're not both in Paris and also in Quebec.
You're in Paris.
∴ You're not in Quebec.

You're not both in Paris and also in Quebec.
You're in Quebec.
∴ You're not in Paris.

Negative parts work the same way; if we affirm one, we can deny the other:
0140

$\sim(\sim A \bullet \sim B)$
$\sim A$
——————————
B

$\sim(A \bullet \sim B)$
A
—————————
B

$\sim(A \bullet \sim B)$
$\sim B$
—————————
$\sim A$

In each case, the second premise affirms (says the same as) one part. And the conclusion denies (says the opposite of) the other part.

If we *deny* one part, we can't draw a conclusion about the other part:

Not both are true.
The first is false.
———————————————
No conclusion.

$\sim(P \bullet Q)$
$\sim P$
———————
nil

You're *not both* in Paris and also in Quebec.
You're not in Paris.
∴ No conclusion.

You may want to conclude "Q"; but maybe "Q" is false too (maybe both parts are false, maybe you're in neither place). To infer with NOT-BOTH, we must *affirm* one part.

Our second I-rule is called "OR," since the larger wff has to have this form:

<table>
<tr><td colspan="2" align="center">OR</td></tr>
<tr><td align="center">(P ∨ Q)
~P
——————
Q</td><td align="center">(P ∨ Q)
~Q
——————
P</td></tr>
<tr><td colspan="2" align="center">deny one part</td></tr>
<tr><td colspan="2" align="center">At least one is true, this one isn't, so the other one is.</td></tr>
</table>

To infer with OR, we must *deny* one part:

At least one hand (left or right) has candy.
The left hand doesn't.
∴ The right hand does.

At least one hand (left or right) has candy.
The right hand doesn't.
∴ The left hand does.

Negative parts work the same; if we deny one part, we can affirm the other:

(~A ∨ ~B)
A
—————————
~B

(A ∨ ~B)
~A
————————
~B

(A ∨ ~B)
B
————————
A

In each case, the second premise denies (says the opposite of) one part. And the conclusion affirms (says the same as) the other part.

If we *affirm* one part, we can't draw a conclusion about the other part:

At least one is true.
The first is true.
–––––––––––––––
No conclusion.

(L ∨ R)
L
––––––
nil

At least one hand (left or right) has candy.
The left hand has candy.
∴ No conclusion. 0141

You may want to conclude "~R"; but maybe "R" is true (maybe both parts are true, maybe both hands have candy). To infer with OR, we must *deny* one part.

Our final I-rule is called "IF-THEN." The first form here is ***modus ponens*** (Latin for "affirming mode") and the second is ***modus tollens*** ("denying mode"):

IF-THEN	
(P ⊃ Q) P –––––– Q	(P ⊃ Q) ~Q –––––– ~P
affirm 1st or deny 2nd	
"IF-THEN, affirm the first, so affirm the second." "IF-THEN, deny the second, so deny the first."	

To infer with IF-THEN, we must *affirm the first* part or *deny the second* part:

If you're a dog, then you're an animal.
You're a dog.
∴ You're an animal.

(D ⊃ A)
D
––––––

A

If you're a dog, then you're an animal.
You're not an animal.
∴ You're not a dog.

(D ⊃ A)
~A
——————
~D

Negative parts work the same. If we affirm the first, we can affirm the second:

(~A ⊃ ~B)
~A
—————————
~B

(A ⊃ ~B)
A
————————
~B

(~A ⊃ B)
~A
————————
B

And if we deny the second, we can deny the first:

(~A ⊃ ~B)
B
—————————
A

(A ⊃ ~B)
B
————————
~A

(~A ⊃ B)
~B
————————
A

If we *deny the first* part or *affirm the second*, we can't conclude anything about the other part:

If you're a dog, then you're an animal.
You're not a dog.
∴ No conclusion.

$$\begin{array}{l} (D \supset A) \\ \underline{\sim D \qquad} \\ \text{nil} \end{array}$$

If you're a dog, then you're an animal.
You're an animal.
∴ No conclusion.

$$\begin{array}{l} (D \supset A) \\ \underline{A \qquad} \\ \text{nil} \end{array}$$

"You're not an animal" doesn't follow in the first case, since you could be a cat. "You're a dog" doesn't follow in the second case, since again you could be a cat. To *infer with an if-then, we need the first part true or the second part false*.

In using I-rules, determine the larger wff's form and apply its rule: 0142

NOT-BOTH	OR	IF-THEN
$\sim(P \bullet Q)$ P ------ $\sim Q$ $\sim(P \bullet Q)$ Q ------ $\sim P$ affirm one part	$(P \vee Q)$ $\sim P$ ------ Q $(P \vee Q)$ $\sim Q$ ------ P deny one part	$(P \supset Q)$ P ------ Q $(P \supset Q)$ $\sim Q$ ------ $\sim P$ affirm 1st or deny 2nd
"NOT-BOTH are true, this one is, so the other one isn't."	"At least one is true, this one isn't, so the other one is."	"IF-THEN, affirm the first, so affirm the second." "IF-THEN, deny the second, so deny the first."

Again, say the little slogan to yourself as you derive the conclusion. (This is much less confusing than saying the individual formulas.)

6.11a Exercise: LogiCola F (IE & IH)

Draw any simple conclusions (a letter or its negation) that follow from these premises. If nothing follows, leave blank.

(∼Q ∨ ∼M) Q ————————
(∼Q ∨ ∼M) Q ———————— ∼M "At least one is true, this one isn't, so the other one is." (OR)

1. ∼(W • T)
   W
   ———————

2. (S ∨ L)
   S
   —————

3. (H ⊃ ∼B)
   H
   ———————

4. (X ⊃ E)
   E
   —————

5. ∼(B • S)
   ∼S
   ——————

6. (∼Y ⊃ K)
   Y
   ———————

7. (K ∨ ∼R)
   R
   ———————

8. ∼(∼S • W)
   ∼W
   ————————

9. $(U \supset G)$
$\underline{U}$
____

10. $(\sim I \vee K)$
$\underline{K}$
____

11. $(C \supset \sim V)$
$\underline{\sim C}$
____

12. $(\sim N \vee \sim A)$
$\underline{A}$
____

13. $\sim(V \bullet H)$
$\underline{\sim V}$
____

14. $(\sim A \supset \sim E)$
$\underline{\sim E}$
____

15. $\sim(\sim F \bullet \sim O)$
$\underline{\sim O}$
____

16. $(Y \vee \sim C)$
$\underline{\sim C}$
____

17. $(\sim L \supset M)$
$\underline{\sim M}$
____

18. $(\sim M \vee \sim B)$
$\underline{\sim M}$
____

19. $\sim(\sim F \bullet \sim Q)$
$\underline{F}$
____

20. $\sim(A \bullet \sim Y)$
$\underline{A}$
____

6.12 Mixing S- and I-rules

Our next exercise mixes S- and I-rule inferences. Use S-rules (the first group below) to simplify *one* premise and I-rules (the second group) to infer from *two* premises:

AND	$\dfrac{(P \bullet Q)}{P,\ Q}$
NOR	$\dfrac{\sim(P \vee Q)}{\sim P,\ \sim Q}$
NIF	$\dfrac{\sim(P \supset Q)}{P,\ \sim Q}$

NOT-BOTH	
$\dfrac{\begin{array}{c}\sim(P \bullet Q)\\ P\end{array}}{\sim Q}$	$\dfrac{\begin{array}{c}\sim(P \bullet Q)\\ Q\end{array}}{\sim P}$
affirm one part	
OR	
$\dfrac{\begin{array}{c}(P \vee Q)\\ \sim P\end{array}}{Q}$	$\dfrac{\begin{array}{c}(P \vee Q)\\ \sim Q\end{array}}{P}$
deny one part	
IF-THEN	
$\dfrac{\begin{array}{c}(P \supset Q)\\ P\end{array}}{Q}$	$\dfrac{\begin{array}{c}(P \supset Q)\\ \sim Q\end{array}}{\sim P}$
affirm 1st or deny 2nd	

In using these rules, focus on the (larger) wff's form. Simplify AND, NOR, and NIF. Infer from NOT-BOTH (with one part true), OR (with one part false), or IF-THEN (with the first part true or the second part false).

6.12a Exercise: LogiCola F (CE & CH)

Draw any simple conclusions (a letter or its negation) that follow from these premises. If nothing follows, leave blank.

(A ⊃ ~B) ~A ––––––––
(no conclusion) "IF-THEN, need first part true or second part false."

1. ~(U • T)
T
–––––––

2. ~(~B ∨ C)
–––––––––

3. (X ⊃ F)
~X
––––––

4. (~S ∨ T)
––––––––

5. (P • ~Q)
–––––––

6. (~I ⊃ ~N)
N
–––––––––

7. (D ∨ ~J)
D
––––––––

8. ~(L • M)
––––––––

9. ~(~C ⊃ D)
–––––––––

10. ~(~R • A)
~R
–––––––––

11. ~(M ∨ ~I)
––––––––––

12. ~(R • ~G)
~G
––––––––––

13. (~L • S)
––––––––

14. (~L ∨ ~T)
L
––––––––––

15. (A ⊃ ~B)
–––––––––

16. ~(W • ~X)
~W
––––––––––

0144

6.13 Extended inferences

S- and I-rules can work on larger formulas too. Suppose you meet a big AND, "((C ≡ D) • (E ⊃ F))." Visualize it as having two parts – and derive both:

((C ≡ D) • (E ⊃ F))
–––––––––––––––––
(C ≡ D), (E ⊃ F)

Say to yourself "AND statement, so both parts are true." Or suppose you meet a big NOR, "~(~A ∨ (B • ~C))." Visualize it as having two parts – and derive the *opposite* of each:

~(~A ∨ (B • ~C))
–––––––––––––––––
A, ~(B • ~C))

Say to yourself "NOT-EITHER is true, so both parts are false." Or suppose you meet a big NIF, "~((C • D) ⊃ (E ⊃ F))." Again, visualize it as having two parts and say "FALSE IF-THEN, so first part true, second part false":

~((C • D) ⊃ (E ⊃ F))
–––––––––––––––––

(C • D), ~(E ⊃ F)

Focus on a complex wff's FORM; we can simplify an AND, NOR, or NIF.

I-rules require two wffs; the larger wff's FORM tells us what further wff we need to complete the inference. Suppose you meet a big NOT-BOTH statement, "~((A ≡ B) • (C • (D ∨ F)))." You can infer with it if you have one part true:

~((A ≡ B) • (C • (D ∨ F)))
(A ≡ B)
––––––––––––––––––––––
~(C • (D ∨ F))

~((A ≡ B) • (C • (D ∨ F)))
(C • (D ∨ F))
––––––––––––––––––––––
~(A ≡ B)

Say to yourself, "NOT-BOTH are true, this one is, so the other one isn't." Or suppose you meet a big OR, "(~A ∨ (B • ~C))." You can infer with it if you have one part false:

(~A ∨ (B • ~C))
A
–––––––––––––––
(B • ~C)

(~A ∨ (B • ~C))
~(B • ~C)
–––––––––––––––
~A

Say to yourself "At least one is true, this one isn't, so the other one is." Or suppose you meet a big IF-THEN, "((C • D) ⊃ (E ⊃ F))." You can infer with it if you have the first part true or the second part false: 0145

((C • D) ⊃ (E ⊃ F))
(C • D)
–––––––––––––––––
(E ⊃ F)

((C • D) ⊃ (E ⊃ F))
~(E ⊃ F)
–––––––––––––––––
~(C • D)

Say to yourself "IF-THEN, affirm the first, so affirm the second" or "IF-THEN, deny the second, so deny the first."

6.14 Logic and computers

Digital computers were developed using ideas from propositional logic. The key insight is that electrical devices can simulate logic formulas.

Computers represent "1" and "0" by different physical states; "1" might be a positive voltage and "0" a zero voltage. An *and-gate* would then be a physical device with two inputs and one output, where the output has a positive voltage if and only if *both* inputs have positive voltages:

An *or-gate* would be similar, except that the output has a positive voltage if and only if *at least one* input has a positive voltage. For any formula, we can construct an input-output device (a *logic gate*) that mimics that formula.

A computer basically converts input information into 1's and 0's, manipulates these by logic gates and memory devices, and converts the resulting 1's and 0's back into a useful output. So propositional logic is central to computers. One of my logic teachers at the University of Michigan, Art Burks, was part of the team in the 1940s that produced the ENIAC, the first large-scale electronic computer. So propositional logic had a key role in moving us into the computer age.

7 Propositional Proofs

Formal proofs are a convenient and powerful way to test arguments. They also help develop our reasoning skills. From now on, formal proofs will be our main method of testing arguments.

7.1 Easier proofs

A formal proof breaks an argument into a series of small steps. We'll use an *indirect proof strategy*, whereby we first assume the opposite of what we want to prove. You may remember such proofs from high-school geometry; to prove that two angles are equal, assume that they *aren't* equal – and then show that this is impossible, because it leads to a contradiction. Similarly, to prove that the butler committed the murder, assume that he *didn't* do it – and then show that this is impossible, because it leads to a contradiction.

Here's an English analog of a formal proof. Suppose we know premises 1 to 4 and want to prove from them that the butler committed the murder:

1 The only people in the mansion were the butler and the maid.
2 If the only people in the mansion were the butler and the maid, then the butler or the maid did it.
3 If the maid did it, then she had a motive.
4 The maid didn't have a motive.
∴ The butler did it.

1 T
2 (T ⊃ (B ∨ M))
3 (M ⊃ H)
4 ∼H
∴ B

First assume that *the butler didn't do it* (∼B). From 1 and 2, conclude that *the butler or the maid did it* (B ∨ M). From 3 and 4, conclude that *the maid didn't do it* (∼M). From these last two, conclude that *the butler did it* (B). This contradicts our assumption, that the butler didn't do it, which is then shown to be false; so therefore, given premises 1 to 4, the butler did it. So the butler is guilty – throw him in jail!

A formal proof is like this, but in symbols. For now, we'll use a three-step strategy: (1) START (assume the conclusion's opposite), (2) S&I (derive further lines using S- and I-rules until we get a contradiction), and (3) RAA (derive the original conclusion). We START this way with our butler argument: 0147

```
1   T
2   (T ⊃ (B ∨ M))
3   (M ⊃ H)
4   ~H
[ ∴ B
5   asm: ~B
```

In the START step here, we block off the conclusion "B" (which reminds us not to use it in deriving further lines) and add "asm:" (for "assume") followed by its simpler contradictory, "~B."

We begin the S&I step by glancing at the **complex wffs** (any wffs longer than a single letter or its negation) and noticing their forms; here the complex wffs are 2 and 3, both IF-THENs. Recall that AND, NOR, and NIF simplify using S-rules, while NOT-BOTH, OR, and IF-THEN can infer using I-rules, if certain extra wffs are available. Since the complex lines here, 2 and 3, are IF-THENs, we can infer with each if we have the first part true or the second part false – which we *do* have (2 has 1, and 3 has 4). So we derive further formulas:

```
  1   T
* 2   (T ⊃ (B ∨ M))
* 3   (M ⊃ H)
  4   ~H
  [ ∴ B
  5   asm: ~B
* 6   ∴ (B ∨ M)    {from 1 and 2}
  7   ∴ M    {from 5 and 6}
  8   ∴ H    {from 3 and 7}
```

Lines 1 and 2 give us line 6 by an IF-THEN rule: "IF-THEN, affirm the first, so affirm the second." Likewise, lines 5 and 6 give us line 7 by the OR rule: "At least one is true, this one isn't, so the other one is." Finally, lines 3 and 7 give us line 8 by an IF-THEN rule: "IF-THEN, affirm the first, so affirm the second." And so we get a contradiction between line 4 and 8 ("~H" and "H").

Here we starred lines 2, 3, and 6 when we used them to derive further formulas. Starring a line tells us that we've used it and so it can be somewhat ignored as we try to derive further steps. I'll talk more about this later.

Once we get a contradiction, as between lines 4 and 8 above, we finish the proof using RAA (*reductio ad absurdum*, reduction to absurdity), which roughly says that an assumption that leads to a contradiction is thereby wrong, and so we can conclude the opposite – which is our original conclusion. At the same time, we block off the lines from the last assumption on down to show that they can't be used in deriving further lines. This finishes our first formal proof: 0148

1 T **Valid**
* 2 (T ⊃ (B ∨ M))
* 3 (M ⊃ H)
4 **~H**
[∴ B
5 ┌ asm: ~B
* 6 │ ∴ (B ∨ M) {from 1 and 2}
7 │ ∴ M {from 5 and 6}
8 └ **∴ H {from 3 and 7}**
9 ∴ B {from 5; 4 contradicts 8}

Now that we've seen a complete proof, we need to firm up the details.

(1) START: Start a proof by blocking off the original conclusion (blocking off tells us to ignore a line for the rest of the proof) and assuming its simpler *contradictory*. Two wffs are **contradictories** if they are exactly alike except that one starts with an additional "~." So if our conclusion is "A," then assume "~A"; but if our conclusion is "~A," then assume "A." And if our conclusion is "(A ⊃ B)," then assume "~(A ⊃ B)." Always add or subtract an initial squiggle to the original conclusion.

(2) S&I: Derive further lines using S- and I-rules until there's a contradiction. Focus on complex wffs that aren't starred or blocked off. Note the forms of these wffs: AND, NOR, and NIF can be simplified, while NOT-BOTH, OR, and IF-THEN can infer if certain other wffs are available. In our sample proof, our first inference has to involve lines 2 or 3, both IF-THENs and the only complex wffs. Often we can do a proof in various ways; so instead of deriving "H" in line 8, we could use 3 and 4 to get "~M," which would contradict 7.

We starred lines 2, 3, and 6. Here are the starring rules – with examples:

Star any wff simplified using an S-rule.

* (A • B)
————————
∴ A
∴ B

Star the *longer* wff used in an I-rule inference.

* (A ⊃ B)
A
————————
∴ B

Starred lines are redundant, since shorter lines have the same information. When you do a proof, focus on *complex wffs that aren't starred or blocked off* and what can be derived from them.[1] While starring is optional, it simplifies your work because it leads you to ignore lines that won't help to derive further formulas. 0149

In the S&I part, we'll use these old S- and I-rules (these and the three new rules hold regardless of what pairs of contradictory wffs replace "P" / "∼P" and "Q" / "∼Q"):

AND	(P • Q) –––––– P, Q
NOR	∼(P ∨ Q) –––––––– ∼P, ∼Q
NIF	∼(P ⊃ Q) –––––––– P, ∼Q

NOT-BOTH	
∼(P • Q) P –––––– ∼Q	∼(P • Q) Q –––––– ∼P
OR	
(P ∨ Q) ∼P –––––– Q	(P ∨ Q) ∼Q –––––– P
IF-THEN	
(P ⊃ Q) P –––––– Q	(P ⊃ Q) ∼Q –––––– ∼P

[1] Once you've starred a complex wff, it's pointless to again use an S-rule on it or to again use it as the *longer* wff in an I-rule inference. So we can focus on complex wffs that aren't starred or blocked off. But it may be useful to use a starred wff as the *smaller* wff in an I-rule inference. Suppose you starred "(A • B)" when you simplified it. If you have an *unstarred* "((A • B) ⊃ C)," feel free to combine it with the starred "(A • B)" to derive "C."

And we'll add three new S-rules – NN, IFF, and NIFF – which we won't use much:

NN

$$\frac{\sim\sim P}{P}$$

NN (NOT-NOT, double negation) eliminates "~~" from the beginning of a wff.

IFF

$$\frac{(P \equiv Q)}{(P \supset Q),\ (Q \supset P)}$$

IFF breaks a biconditional into two conditionals.

NIFF

$$\frac{\sim(P \equiv Q)}{(P \vee Q),\ \sim(P \bullet Q)}$$

NIFF[1] breaks up the denial of a biconditional; since "(P ≡ Q)" says that P and Q have the *same* truth value, "~(P ≡ Q)" says that P and Q have *different* truth values – so one or the other is true, but not both.[2]

In applying S- and I-rules, look for lines of these forms to simplify:

AND NOR NIF
NN IFF NIFF

or a pair of lines to infer from:

[1] To avoid confusion, pronounce "NIFF" as "knife" and "IFF" with a long "i" to rhyme with this.

[2] The S-rules also work in the other direction (so "(A • B)" follows from "A" and "B"); but our proofs standardly use S-rules only to simplify. The LogiCola software lets you use two further rules: (1) Given "(A ≡ B)": if you have one side true, you can get the other true – and if you have one side false, you can get the other false. (2) Given "~(A ≡ B)": if you have one side true, you can get the other false – and if you have one side false, you can get the other true.

NOT-BOTH (with one part true)
OR (with one part false)
IF-THEN (with part-1 true or part-2 false)

Note that there's a rule for each of the nine possible complex wff forms.

Here's another summary of the S- and I-rules (here "→" means we can infer whole lines from left to right): 0150

S-rules *(Simplifying)*

AND, NOR, NIF, NN, IFF, NIFF

$$(P \bullet Q) \rightarrow P, Q$$
$$\sim(P \vee Q) \rightarrow \sim P, \sim Q$$
$$\sim(P \supset Q) \rightarrow P, \sim Q$$
$$\sim\sim P \rightarrow P$$
$$(P \equiv Q) \rightarrow (P \supset Q), (Q \supset P)$$
$$\sim(P \equiv Q) \rightarrow (P \vee Q), \sim(P \bullet Q)$$

I-rules *(Inferring)*

NOT-BOTH, OR, IF-THEN

$$\sim(P \bullet Q), P \rightarrow \sim Q$$
$$\sim(P \bullet Q), Q \rightarrow \sim P$$
$$(P \vee Q), \sim P \rightarrow Q$$
$$(P \vee Q), \sim Q \rightarrow P$$
$$(P \supset Q), P \rightarrow Q$$
$$(P \supset Q), \sim Q \rightarrow \sim P$$

Read "(P • Q) → P, Q" as "from '(P • Q)' one may derive 'P' and also 'Q.'" As you learn formal proofs, it's good to practice the S- and I-rules.

(3) RAA: Rule RAA says roughly that an assumption is false if it leads to *contradictory wffs* (a pair, like "H" and "∼H," that's identical except that one starts with an additional squiggle). The contradictory wffs may occur anywhere in the proof (as premises, assumptions, or derived lines), as long as neither is blocked off. Here's a more precise formulation of RAA:

> **RAA**: Suppose some pair of not-blocked-off lines has contradictory wffs. Then block off all the lines from the last not-blocked-off assumption on down and infer a line consisting in "∴" followed by a contradictory of that assumption.

Blocking off forbids deriving further lines using the assumption (which now is shown to be false). This is important later, with multiple-assumption proofs.

Here are some key definitions about formal proofs:

- A **premise** is a line consisting of a wff by itself (with no "asm:" or "∴").
- An **assumption** is a line consisting of "asm:" and then a wff.
- A **derived line** is a line consisting of "∴" and then a wff.
- A **formal proof** is a vertical sequence of zero or more premises followed by one or more assumptions or derived lines, where each derived line follows from previously not-blocked-off lines by one of the S- and I-rules listed above or by RAA, and each assumption is blocked off using RAA.

By the last definition, the stars, line numbers, blocked off original conclusion, and justifications aren't strictly part of the proof. Instead, these are unofficial helps – and some people skip them. On the other hand, some people like to mention the inference rule (like "AND" or "IF-THEN") in the justifications; so then a justification might say "{from 1 and 2 using IF-THEN}." If you're taking a logic course, follow your teacher's directives about such matters.

A wff is a **theorem** if it's provable from zero premises. Here's a premiseless proof (it's valid because the conclusion is a logically necessary truth): 0151

[∴ ((A • B) ⊃ A)   **Valid**
* 1 asm: ~((A • B) ⊃ A)
2 ∴ (A • B)   {from 1}
3 **∴ ~A   {from 1}**
4 **∴ A   {from 2}**
5 ∴ ((A • B) ⊃ A)   {from 1; 3 contradicts 4}

Again, our proof strategy has three steps. (1) START: Block off the conclusion and assume its contradictory (line 1). (2) S&I: Derive lines 2 to 4 and get a contradiction. (3) RAA: Use RAA to finish the proof (line 5). Our proof strategy gets more complex later, with invalid arguments and multiple assumptions.

A formal proof, as we defined it, must use the specified S- and I-rules or RAA to derive further lines. We can't just use any intuitive inferences that we think will work (although advanced users sometimes take such shortcuts[1]).

[1] An example of a shortcut is to infer "C" immediately from previous lines "(A ⊃ (B ⊃ C))" and "A" and "B" in a single step. Don't take such shortcuts unless your teacher allows them and you're very sure that your formula validly follows, even though not licensed by our rules. I suggest that you thoroughly master our normal strategy before taking shortcuts; otherwise, your reliance on shortcuts may lead to steps that don't validly follow and may prevent you from learning our proof procedure (which will always work if you do it right). LogiCola doesn't allow shortcuts.

Why not add further inference rules to our system, since this would shorten some proofs?

There can be legitimate variations in how to do proofs. So one person might always simplify "(A • B)" into the two parts, "A" first and then "B." Another might derive "B" first and then "A." Yet another person might derive just the part needed to get a contradiction. All three approaches are fine and allowed by the LogiCola computer proof exercises.

LogiCola proofs begin by giving you a randomly generated problem (there are many millions of possible problems). You keep giving the next line until the problem is done. You can vary the kind of problem: Easier / Harder / Mixed – and Valid / Invalid / Combined. As you begin, turn on training wheels; this gives you suggestions about what to do next – but these suggestions disappear as you make progress in the exercise. You can also have the program automatically star lines that you've used, or you can choose to star yourself (but you don't lose points for getting these wrong). You can click (or touch) an arrow at the top to give you the next line or to finish the problem (but without getting credit for the problem); some students use these arrows to step through sample proofs before starting them on their own. You can click (or touch) a previous line to copy it into the answer space, so you can then modify it to give your next line. If you're new to proofs, I suggest you read the LogiCola help-section on proofs.

7.1a Exercise: LogiCola F (TE & TH) and GEV

Prove each of these arguments to be valid (all are valid). 0152

(A ∨ B)
∴ (~A ⊃ B)

* 1 (A ∨ B) **Valid**
[∴ (~A ⊃ B)
* 2 ┌ asm: ~(~A ⊃ B)
3 │ ∴ ~A {from 2}
4 │ **∴ ~B {from 2}**
5 └ **∴ B {from 1 and 3}**
6 ∴ (~A ⊃ B) {from 2; 4 contradicts 5}

1. (A ⊃ B)
∴ (~B ⊃ ~A)

2. A
∴ (A ∨ B)

The downside is that this would also make our system harder to learn. Our proof system was designed in a practical way, to produce reasonably short proofs and yet be easy to learn and use.

3. (A ⊃ B)
(∼A ⊃ B)
∴ B

4. ((A ∨ B) ⊃ C)
∴ (∼C ⊃ ∼B)

5. (A ∨ B)
(A ⊃ C)
(B ⊃ D)
∴ (C ∨ D)

6. (A ⊃ B)
(B ⊃ C)
∴ (A ⊃ C)

7. (A ≡ B)
∴ (A ⊃ (A • B))

8. ∼(A ∨ B)
(C ∨ B)
∼(D • C)
∴ ∼D

9. (A ⊃ B)
∼B
∴ (A ≡ B)

10. (A ⊃ (B ⊃ C))
∴ ((A • B) ⊃ C)

7.1b Exercise: LogiCola F (TE & TH) and GEV

First appraise intuitively. Then translate into logic (using the letters given) and prove to be valid (all are valid).

1. If Heather saw the butler putting the tablet into the drink and the tablet was poison, then the butler killed the deceased.
Heather saw the butler putting the tablet into the drink.
∴ If the tablet was poison, then the butler killed the deceased. [Use H, T, and B.]

2. If we had an absolute proof of God's existence, then our will would be irresistibly attracted to do right.
If our will were irresistibly attracted to do right, then we'd have no free will.
∴ If we have free will, then we have no absolute proof of God's existence. [Use P, I, and F; from Immanuel Kant and John Hick, who used it to explain why God doesn't make his existence more evident.]

3. If racism is clearly wrong, then either it's factually clear that all races have equal abilities or it's morally clear that similar interests of all beings ought to be given equal consideration.
It's not factually clear that all races have equal abilities.
If it's morally clear that similar interests of all beings ought to be given equal consideration, then similar interests of animals and humans ought to be given equal consideration.
∴ If racism is clearly wrong, then similar interests of animals and humans ought to be given equal consideration. [Use W, F, M, and A. This argument is from Peter Singer, who fathered the animal liberation movement.] 0153

4. The universe is orderly (like a watch that follows complex laws).
Most orderly things we've examined have intelligent designers.
We've examined a large and varied group of orderly things.
If most orderly things we've examined have intelligent designers and we've examined a large and varied group of orderly things, then probably most orderly things have intelligent designers.
If the universe is orderly and probably most orderly things have intelligent designers, then the universe probably has an intelligent designer.
∴ The universe probably has an intelligent designer. [Use U, M, W, P, and D. This is a form of the argument from design for the existence of God.]

5. If God doesn't want to prevent evil, then he isn't all good.
If God isn't able to prevent evil, then he isn't all powerful.
Either God doesn't want to prevent evil, or he isn't able.
∴ Either God isn't all powerful, or he isn't all good. [Use W, G, A, and P. This form of the problem-of-evil argument is from the ancient Greek Empiricus.]

6. If Genesis gives the literal facts, then birds were created before humans. (Genesis 1:20–26)
If Genesis gives the literal facts, then birds weren't created before humans. (2:5–20)
∴ Genesis doesn't give the literal facts. [Use L and B. Origen, an early Christian thinker, gave similar textual arguments against taking Genesis literally.]

7. The world had a beginning in time.
If the world had a beginning in time, there was a cause for the world's beginning.
If there was a cause for the world's beginning, a personal being caused the world.
∴ A personal being caused the world. [Use B, C, and P. This "Kalam argument" for the existence of God is from William Craig and James Moreland; they defend premise 1 by various considerations, including the Big Bang theory, the law of entropy, and the impossibility of an actual infinite.]

8. If the world had a beginning in time and it didn't just pop into existence without any cause, then the world was caused by God.
If the world was caused by God, then there is a God.
There is no God.
∴ Either the world had no beginning in time, or it just popped into existence without any cause. [Use B, P, C, and G; from J. L. Mackie, who based his "There is no God" premise on the problem-of-evil argument.]

9. Closed systems tend toward greater entropy (a more randomly uniform distribution of energy). (This is the second law of thermodynamics.)
If closed systems tend toward greater entropy and the world has existed through endless time, then the world would have achieved almost complete entropy (for example, everything would be about the same temperature).
The world has not achieved almost complete entropy.
If the world hasn't existed through endless time, then the world had a beginning in time.
∴ The world had a beginning in time. [Use G, E, C, and B; from William Craig and James Moreland.] 0154

10. If time stretches back infinitely, then today wouldn't have been reached.
If today wouldn't have been reached, then today wouldn't exist.
Today exists.
If time doesn't stretch back infinitely, then there was a first moment of time.
∴ There was a first moment of time. [I, R, T, F]

11. If there are already laws preventing discrimination against women, then if the Equal Rights Amendment (ERA) would rob women of many current privileges then it is the case both that passage of the ERA would be against women's interests and that women ought to work for its defeat.
The ERA would rob women of many current privileges (like draft exemption).
∴ If there are already laws preventing discrimination against women, then women ought to work for the defeat of the ERA. [L, R, A, W]

12. If women ought never to be discriminated against, then we should pass current laws against discrimination and block future discriminatory laws against women.
The only way to block future discriminatory laws against women is to pass an Equal Rights Amendment (ERA).
If we should block future discriminatory laws against women and the only way to do this is to pass an ERA, then we ought to pass an ERA.
∴ If women ought never to be discriminated against, then we ought to pass an ERA. [N, C, F, O, E]

13. If the claim that knowledge-is-impossible is true, then we understand the word "know" but there are no cases of knowledge.
If we understand the word "know," then the meaning of "know" comes either from a verbal definition or from experienced examples of knowledge.
If the meaning of "know" comes from a verbal definition, then there's an agreed-upon definition of "know."
There's no agreed-upon definition of "know."
If the meaning of "know" comes from experienced examples of knowledge, then there are cases of knowledge.
∴ The claim that knowledge-is-impossible is false. [Use I, U, C, D, E, and A. This is a form of the paradigm-case argument.]

14. If *p* is the greatest prime, then *n* (we may stipulate) is one plus the product of all the primes less than *p*.
If *n* is one plus the product of all the primes less than *p*, then either *n* is prime or else *n* isn't prime but has prime factors greater than *p*.
If *n* is prime, then *p* isn't the greatest prime.
If *n* has prime factors greater than *p*, then *p* isn't the greatest prime.
∴ *p* isn't the greatest prime. [Use G, N, P, and F. This proof that there's no greatest prime number is from the ancient Greek mathematician Euclid.]

7.2 Easier refutations

This example shows how our proof strategy works with an invalid argument:
0155

The only people in the mansion were the butler and the maid.
If the only people in the mansion were the butler and the maid, then the butler or the maid did it.
If the maid did it, then she had a motive.
∴ The butler did it.

T
(T ⊃ (B ∨ M))

(M ⊃ H)
∴ B

The butler's lawyer could object: "Yes, the only people in the mansion were the butler and the maid, and so one of them did the killing. But maybe the maid had a motive and did it, instead of the butler. The known facts are consistent with this possibility and so don't show that the butler did it." This is a **refutation** – a set of possible truth conditions making the premises all true and conclusion false. A refutation shows that the argument is invalid.

If we try to prove this invalid argument, we'll assume the conclusion's opposite and then use S- and I-rules to derive whatever we can:

  1 T
* 2 (T ⊃ (B ∨ M))
* 3 (M ⊃ H)
  [∴ B
  4 asm: ~B
* 5 ∴ (B ∨ M) {from 1 and 2}
  6 ∴ M {from 4 and 5}
  7 ∴ H {from 3 and 6}

We can derive no contradiction. So we instead construct a refutation box – which contains the **simple wffs** (letters or their negations) from not-blocked-off lines (1, 4, 6, and 7) – and we plug its truth values into the original argument:

1 T^1 = 1 **Invalid**
2 (T^1 ⊃ (B^0 ∨ M^1)) = 1
3 (M^1 ⊃ H^1) = 1
[∴ B^0 = 0

T, M, H, ~B

These truth conditions make the premises all true and conclusion false. This shows that the argument is invalid.

With invalid arguments, we don't get a contradiction; instead, we get a refutation. To construct the refutation box, take the simple wffs (letters or their negation) from not-blocked-off lines and put them in a box (their order doesn't matter). Our box also could be written in either of these two ways:

T = 1, M = 1, H = 1, B = 0
T^1, M^1, H^1, B^0

Then plug the truth values into the original argument. If the refutation box has a letter by itself (like “T” or “M”), then mark that letter *true* (“1”) in the
0156 argument; if it has the negation of a letter (like “∼B”), then mark that letter *false* (“0”); any letters that don’t occur in the box are *unknown* (“?” – the refutation may still work). Then see if these values make the premises all true and conclusion false; if they do, then that shows that the argument is invalid.

If we *don’t* get the premises all true and conclusion false, then we did something wrong. The faulty line (a premise that’s false or unknown, or a conclusion that’s true or unknown) is the problem’s source; maybe we derived something from it wrongly, or didn’t derive something we should have derived. So our strategy tells us if something goes wrong and where to look to fix the problem.

Let me summarize. Suppose we want to show that, given certain premises, the butler must be guilty. We assume that he’s innocent and try to show that this leads to a contradiction. If we get a contradiction, then his innocence is impossible and so he must be guilty. But if we get no contradiction, then we may be able to show how the premises could be true while yet he is innocent, thus showing that the argument against him is invalid.

Here’s another invalid argument and its refutation:

1 $(A^0 \supset B^1) = 1$ **Invalid**
* 2 $(C^0 \vee B^1) = 1$
[∴ $(C^0 \vee A^0) = 0$
* 3 asm: ∼(C ∨ A)
4 ∴ ∼C {from 3}
5 ∴ ∼A {from 3}
6 ∴ B {from 2 and 4}

B, ∼A, ∼C

We get nothing from “(A ⊃ B)” in line 1, since we’d need “A” true or “B” false. So we’ve derived all we can. Since we have no contradiction, we construct a refutation box. We plug the values into the argument and get the premises all true and conclusion false. This shows that the argument is invalid.

Our proof strategy so far looks like this (we’ll add another step later):

1. START: Block off the conclusion and add “asm:” followed by the conclusion’s simpler contradictory.

2. S&I: Go through the complex wffs that aren’t starred or blocked off and use these to derive new wffs using S- and I-rules. Star any wff you simplify using an S-rule, or the longer wff used in an I-rule inference. If you get a contradiction, then go to RAA (step 3). If you can’t derive anything further and yet have no contradic-

tion, then go to REFUTE (step 4).

3. RAA: Apply the RAA rule. You've proved the argument valid.

4. REFUTE: Construct a refutation box containing any simple wffs (letters or their negation) that aren't blocked off. In the original argument, mark each letter "1" or "0" or "?" depending on whether the box has the letter or its negation or neither. If these truth conditions make the premises all true and conclusion false, then this shows the argument to be invalid. 0157

7.2a Exercise: LogiCola GEI

Prove each of these arguments to be invalid (all are invalid).

(A ⊃ B)
∴ (B ⊃ A)

1 (A^0 ⊃ B^1) = 1 **Invalid**
[∴ (B^1 ⊃ A^0) = 0
* 2 asm: ~(B ⊃ A)
3 ∴ B {from 2}
4 ∴ ~A {from 2}

B, ~A

1. (A ∨ B)
∴ A

2. (A ⊃ B)
(C ⊃ B)
∴ (A ⊃ C)

3. ~(A • ~B)
∴ ~(B • ~A)

4. (A ⊃ (B • C))
(~C ⊃ D)
∴ ((B • ~D) ⊃ A)

5. ((A ⊃ B) ⊃ (C ⊃ D))
(B ⊃ D)
(A ⊃ C)
∴ (A ⊃ D)

6. (A ≡ B)
(C ⊃ B)
∼(C • D)
D
∴ ∼A

7. ((A • B) ⊃ C)
∴ (B ⊃ C)

8. ((A • B) ⊃ C)
((C ∨ D) ⊃ ∼E)
∴ ∼(A • E)

9. ∼(A • B)
(∼A ∨ C)
∴ ∼(C • B)

10. ∼(∼A • ∼B)
∼C
(D ∨ ∼A)
((C • ∼E) ⊃ ∼B)
∼D
∴ ∼E

7.2b Exercise: LogiCola GEC

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation).

1. If the butler shot Jones, then he knew how to use a gun.
If the butler was a former marine, then he knew how to use a gun.
The butler was a former marine.
∴ The butler shot Jones. [Use S, K, and M.]

2. If virtue can be taught, then either there are professional virtue-teachers or there are amateur virtue-teachers.
If there are professional virtue-teachers, then the Sophists can teach their students to be virtuous.
If there are amateur virtue-teachers, then the noblest Athenians can teach their children to be virtuous.
The Sophists can't teach their students to be virtuous and the noblest Athenians (such as the great leader Pericles) can't teach their children to be virtuous.
∴ Virtue can't be taught. [Use V, P, A, S, and N; from Plato's *Meno*.] 0158

3. It would be equally wrong for a sadist (through drugs that would blind you but not hurt your mother) to have blinded you permanently before or after your birth.
If it would be equally wrong for a sadist (through such drugs) to have blinded you permanently before or after your birth, then it's false that one's moral right to equal consideration begins at birth.
If infanticide is wrong and abortion isn't wrong, then one's moral right to equal consideration begins at birth.
Infanticide is wrong.
∴ Abortion is wrong. [Use E, R, I, and A.]

4. If you hold a moral belief and don't act on it, then you're inconsistent.
If you're inconsistent, then you're doing wrong.
∴ If you hold a moral belief and act on it, then you aren't doing wrong. [Use M, A, I, and W. Is the conclusion plausible? What more plausible conclusion follows from these premises?]

5. If Socrates escapes from jail, then he's willing to obey the state only when it pleases him.
If he's willing to obey the state only when it pleases him, then he doesn't really believe what he says and he's inconsistent.
∴ If Socrates really believes what he says, then he won't escape from jail. [Use E, W, R, and I; from Plato's *Crito*. Socrates had been jailed and sentenced to death for teaching philosophy. He discussed with his friends whether he ought to escape from jail instead of suffering the death penalty.]

6. Either Socrates's death will be perpetual sleep, or if the gods are good then his death will be an entry into a better life.
If Socrates's death will be perpetual sleep, then he shouldn't fear death.
If Socrates's death will be an entry into a better life, then he shouldn't fear death.
∴ Socrates shouldn't fear death. [Use P, G, B, and F; from Plato's *Crito* – except for which dropped premise?]

7. If predestination is true, then God causes us to sin.
If God causes us to sin and yet damns sinners to eternal punishment, then God isn't good.
∴ If God is good, then either predestination isn't true or else God doesn't damn sinners to eternal punishment. [Use P, C, D, and G. This attacks the views of the American colonial thinker Jonathan Edwards.]

8. If determinism is true, then we have no free will.
If Heisenberg's interpretation of quantum physics is correct, some events aren't causally necessitated by prior events.
If some events aren't causally necessitated by prior events, determinism is false.
∴ If Heisenberg's interpretation of quantum physics is correct, then we have free will. [D, F, H, E] 0159

9. Government's function is to protect life, liberty, and the pursuit of happiness.
The British colonial government doesn't protect these.
The only way to change it is by revolution.
If government's function is to protect life, liberty, and the pursuit of happiness and the British colonial government doesn't protect these, then the British colonial government ought to be changed.
If the British colonial government ought to be changed and the only way to change it is by revolution, then we ought to have a revolution.
∴ We ought to have a revolution. [Use G, B, O, C, and R. This summarizes the reasoning behind the American Declaration of Independence. Premise 1 was claimed to be self-evident, premises 2 and 3 were backed by historical data, and premises 4 and 5 were implicit conceptual bridge premises.]

10. The apostles' teaching either comes from God or is of human origin.
If it comes from God and we kill the apostles, then we will be fighting God.
If it's of human origin, then it'll collapse of its own accord.
If it'll collapse of its own accord and we kill the apostles, then our killings will be unnecessary.
∴ If we kill the apostles, then either our killings will be unnecessary or we will be fighting God. [Use G, H, K, F, C, and U. This argument, from Rabbi Gamaliel in Acts 5:34–9, is perhaps the most complex reasoning in the Bible.]

11. If materialism (the view that only matter exists) is true, then idealism is false.
If idealism (the view that only minds exist) is true, then materialism is false.
If mental events exist, then materialism is false.
If materialists *think* their theory is true, then mental events exist.
∴ If materialists *think* their theory is true, then idealism is true. [M, I, E, T]

12. If determinism is true and cruelty is wrong, then the universe contains unavoidable wrong actions.
If the universe contains unavoidable wrong actions, then we ought to regret the universe as a whole.
If determinism is true and regretting cruelty is wrong, then the universe contains unavoidable wrong actions.
∴ If determinism is true, then either we ought to regret the universe as a whole (the pessimism option) or else cruelty isn't wrong and regretting cruelty isn't wrong (the "nothing matters" option). [Use D, C, U, O, and R. This sketches the reasoning in William James's "The Dilemma of Determinism." James thought that when we couldn't prove one side or the other to be correct (as on the issue of determinism), it was more rational to pick our beliefs in accord with practical considerations. He argued that these weighed against determinism.]

13. If a belief is proved, then it's worthy of acceptance.
If a belief isn't disproved but is of practical value to our lives, then it's worthy of acceptance.
If a belief is proved, then it's not disproved.
∴ If a belief is proved or is of practical value to our lives, then it's worthy of acceptance. [P, W, D, V] 0160

14. If you're consistent and think that stealing is normally permissible, then you'll consent to the idea of others stealing from you in normal circumstances.
You don't consent to the idea of others stealing from you in normal circumstances.
∴ If you're consistent, then you won't think that stealing is normally permissible. [C, N, Y]

15. If the meaning of a term is always the object it refers to, then the meaning of "Fido" is Fido.
If the meaning of "Fido" is Fido, then if Fido is dead then the meaning of "Fido" is dead.
If the meaning of "Fido" is dead, then "Fido is dead" has no meaning.
"Fido is dead" has meaning.
∴ The meaning of a term isn't always the object it refers to. [Use A, B, F, M, and H; from Ludwig Wittgenstein, except for which dropped premise?]

16. God is all powerful.
If God is all powerful, then he could have created the world in any logically possible way and the world has no necessity.
If the world has no necessity, then we can't know the way the world is by abstract speculation apart from experience.
∴ We can't know the way the world is by abstract speculation apart from experience. [Use A, C, N, and K; from the medieval William of Ockham.]

17. If God changes, then he changes for the worse or for the better.
If he's perfect, then he doesn't change for the worse.
If he changes for the better, then he isn't perfect.
∴ If God is perfect, then he doesn't change. [C, W, B, P]

18. If belief in God has scientific backing, then it's rational.
No conceivable scientific experiment could decide whether there is a God.
If belief in God has scientific backing, then some conceivable scientific experiment could decide whether there is a God.
∴ Belief in God isn't rational. [B, R, D]

19. Every event with finite probability eventually takes place.
If the nations of the world don't get rid of their nuclear weapons, then there's a finite probability that humanity will eventually destroy the world.
If every event with finite probability eventually takes place and there's a finite probability that humanity will eventually destroy the world, then humanity will eventually destroy the world.
∴ Either nations of the world will get rid of their nuclear weapons, or humanity will eventually destroy the world. [E, R, F, H]

20. If the world isn't ultimately absurd, then conscious life will go on forever and the world process will culminate in an eternal personal goal.
If there is no God, then conscious life won't go on forever.
∴ If the world isn't ultimately absurd, then there is a God. [Use A, F, C, and G; from the Jesuit scientist, Pierre Teilhard de Chardin.] 0161

21. If it rained here on this date 500 years ago and there's no way to know whether it rained here on this date 500 years ago, then there are objective truths that we cannot know.
If it didn't rain here on this date 500 years ago and there's no way to know whether it rained here on this date 500 years ago, then there are objective truths that we cannot know.
There's no way to know whether it rained here on this date 500 years ago.
∴ There are objective truths that we cannot know. [R, K, O]

22. If you know that you don't exist, then you don't exist.
If you know that you don't exist, then you know some things.
If you know some things, then you exist.
∴ You exist. [K, E, S]

23. We have an idea of a perfect being.
If we have an idea of a perfect being, then this idea is either from the world or from a perfect being.
If this idea is from a perfect being, then there is a God.
∴ There is a God. [Use I, W, P, and G; from René Descartes, except for which dropped premise?]

24. The distance from A to B can be divided into an infinity of spatial points.
One can cross only one spatial point at a time.
If one can cross only one spatial point at a time, then one can't cross an infinity of spatial points in a finite time.
If the distance from A to B can be divided into an infinity of spatial points and one can't cross an infinity of spatial points in a finite time, then one can't move from A to B in a finite time.
If motion is real, then one can move from A to B in a finite time.
∴ Motion isn't real. [Use D, O, C, M, and R; from the ancient Greek Zeno of Elea, who denied the reality of motion.]

25. If the square root of 2 equals some fraction of positive whole numbers, then (we stipulate) the square root of 2 equals x/y and x/y is simplified as far as it can be.
If the square root of 2 equals x/y, then $2 = x^2/y^2$.
If $2 = x^2/y^2$, then $2y^2 = x^2$.
If $2y^2 = x^2$, then x is even.
If x is even and $2y^2 = x^2$, then y is even.
If x is even and y is even, then x/y isn't simplified as far as it can be.
∴ The square root of 2 doesn't equal some fraction of positive whole numbers. [F, E, S, T, T´, X, Y]

7.3 Harder proofs

Our present proof strategy has four steps: START, S&I, RAA, and REFUTE. Some arguments require a further multiple-assumption ASSUME step. Here's an example: 0162

If the butler was at the party, then he fixed the drinks and poisoned the deceased.
If the butler wasn't at the party, then the detective would have seen him leave the mansion and would have reported this.
The detective didn't report this.
∴ The butler poisoned the deceased.

(A ⊃ (F • P))
(∼A ⊃ (S • R))
∼R
∴ P

START by assuming "∼P":

1 (A ⊃ (F • P))
2 (∼A ⊃ (S • R))
3 ∼R
[∴ P
4 asm: ∼P

Then we're stuck. We can't apply the S- or I-rules or RAA; and we don't have enough simple wffs for a refutation. What can we do? On our newly expanded strategy, when we get stuck we'll make another assumption. We pick a complex wff we haven't used yet (1 or 2), pick left or right side, and assume it or its negation. Here we decide to assume the negation of the left side of line 1:

1 (A ⊃ (F • P))
2 (∼A ⊃ (S • R))
3 ∼R
[∴ P
4 asm: ∼P
5 asm: ∼A {break 1}

We use S- and I-rules to derive further lines; but now we use two stars (one for each assumption). Lines 3 and 8 contradict:

1 (A ⊃ (F • P))
** 2 (∼A ⊃ (S • R))
3 **∼R**
[∴ P
4 asm: ∼P
5 asm: ∼A {break 1}
** 6 ∴ (S • R) {from 2 and 5}
7 ∴ S {from 6}
8 **∴ R {from 6}**

Since we have a contradiction, we (1) block off the lines from the last assumption on down (this tells us not to use these lines, here 5 to 8, as we derive further lines and look for a contradiction), (2) derive the opposite of this last assumption, and (3) erase star strings with more stars than the number of remaining assumptions: 0163

1 (A ⊃ (F • P))
2 (∼A ⊃ (S • R))
3 ∼R
[∴ P
4 asm: ∼P
5 asm: ∼A {break 1}
6 ∴ (S • R) {from 2 and 5}
7 ∴ S {from 6}
8 ∴ R {from 6}
9 ∴ A {from 5; 3 contradicts 8}

Then we use S- and I-rules to derive further lines, and thus we get a second contradiction (lines 4 and 12):

* 1 (A ⊃ (F • P))
2 (∼A ⊃ (S • R))
3 ∼R
[∴ P

```
   4   asm: ~P
   5   ┌ asm: ~A   {break 1}
   6   │ ∴ (S • R)   {from 2 and 5}
   7   │ ∴ S   {from 6}
   8   └ ∴ R   {from 6}
   9   ∴ A   {from 5; 3 contradicts 8}
*  10  ∴ (F • P)   {from 1 and 9}
   11  ∴ F   {from 10}
   12  ∴ P   {from 10}
```

Finally, we apply RAA again, this time on our original assumption:

```
*  1  (A ⊃ (F • P))   Valid
   2  (~A ⊃ (S • R))
   3  ~R
  [ ∴ P
   4 ┌ asm: ~P
   5 │ ┌ asm: ~A   {break 1}
   6 │ │ ∴ (S • R)   {from 2 and 5}
   7 │ │ ∴ S   {from 6}
   8 │ └ ∴ R   {from 6}
   9 │ ∴ A   {from 5; 3 contradicts 8}
*  10 │ ∴ (F • P)   {from 1 and 9}
   11 │ ∴ F   {from 10}
   12 └ ∴ P   {from 10}
   13 ∴ P   {from 4; 4 contradicts 12}
```

To prove the argument valid, we need to get a contradiction for each assumption. We've accomplished this, and our proof is done. 0164

The most difficult part of multiple-assumption proofs is knowing *when* to make another assumption and *what* to assume.

(1) *Make another assumption when you're stuck.* You may get that deep sense of confusion in your gut. More technically, *being stuck* means that you can't apply S- or I-rules further – and yet you can't prove the argument VALID (since you have no contradiction) or INVALID (since you don't have enough simple wffs for a refutation). Don't make additional assumptions too soon; it's *too soon* if you can still apply S- or I-rules or RAA. Always use S- and I-rules and RAA to their limit before resorting to further assumptions.

(2) *When you're stuck, make an assumption that breaks a complex wff.* Look for a complex wff that isn't starred, blocked off, or *broken* (a wff is broken if we already have one side or its negation but not what we need to conclude anything new). This wff will have a NOT-BOTH, OR, or IF-THEN form:

~(A • B)
(A ∨ B)
(A ⊃ B)

Assume either side or its negation. Here we could use any of these:

asm: A
asm: ~A
asm: B
asm: ~B

While any of the four works, our proof will go differently depending on which we use. Suppose we want to break "(A ⊃ B)"; compare what happens if we assume "A" or assume "~A":

(immediate gratification)

(A ⊃ B)
asm: A
∴ B

(delayed gratification)

(A ⊃ B)
asm: ~A
...

In the first case, we assume "A" and use an I-rule on "(A ⊃ B)" to get "B." In the second case, we assume "~A" and get nothing; but we may be able to use an I-rule on "(A ⊃ B)" later, after the "~A" assumption dies (if it does) and we derive "A." *Delayed gratification* tends to produce shorter proofs; it saves an average of one line, with the gain coming on invalid arguments. So sometimes a proof is simpler if you assume one thing rather than another.

Do the same with longer wffs. To break "((A • B) ⊃ (C • D))," make any of these four assumptions:

asm: (A • B)
asm: ~(A • B)
asm: (C • D)
asm: ~(C • D)

Assume one side or its negation. Never assume the denial of a whole line.

Never make an assumption to break a wff that's already *broken*. A wff is *broken* if we already have one side or its negation but not what we need to conclude anything new. So a "(A ⊃ B)" line, for example, is broken if we already have a not-blocked-off line with "~A" or with "B." In such a case, it won't help us to make an assumption to break "(A ⊃ B)." 0165

After making our second assumption, we star the same things as before, but now we use more stars:

Use one star for each live assumption.

Star any wff simplified using an S-rule.

```
** (A • B)
 ────────
   ∴ A
   ∴ B
```

Star the *longer* wff used in an I-rule inference.

```
** (A ⊃ B)
      A
 ────────
   ∴ B
```

A *live assumption* is one that isn't blocked off. So if we have two live assumptions, then we use two stars. And if we have three live assumptions, then we use three stars. As before, starred lines are redundant; when doing a proof, focus on *complex wffs that aren't starred or blocked off* and what can be derived from them. Multiple stars mean "You can ignore this line for now, but you may have to use it later."

When we have multiple live assumptions and find a contradiction:

- block off the lines from the last live assumption on down (these lines are no longer to be used in the proof – since they depend on an assumption that we've concluded to be false);
- derive the opposite of this last assumption; and
- erase star strings with more stars than the number of remaining live assumptions (since the blocked-off lines that make these starred lines redundant are no longer available).

Note the part about erasing star strings with more stars than the number of remaining live assumptions. So if our second assumption dies, leaving us with just one live assumption, then we erase double-stars ("**").

When our last live assumption leads to a contradiction, we've proved the argument to be valid. Valid arguments seldom require more than two assumptions. But if we get stuck again after making a second assumption, then we'll need to make a third assumption.

Our final proof strategy can prove or refute any propositional argument (as we'll show in §15.4):

1. START: Block off the conclusion and add "asm:" followed by the conclusion's simpler contradictory.

2. S&I: Go through the complex wffs that aren't starred or blocked off and use these to derive new wffs using S- and I-rules. Star (with one star for each live assumption) any wff you simplify using an S-rule, or the longer wff used in an I-rule inference. If you get a contradiction, then go to RAA (step 3). If you can't derive anything further but there is a complex wff that isn't starred or blocked off or broken, then go to ASSUME (step 4). If you can't derive anything further and every complex wff is starred or blocked off or broken, then go to REFUTE (step 5).
0166

3. RAA: Apply the RAA rule. If all assumptions are now blocked off, you've proved the argument valid. Otherwise, erase star strings having more stars than the number of live assumptions and return to step 2.

4. ASSUME: Pick a complex wff that isn't starred or blocked off or broken. This wff will have one of these forms: "∼(A • B)," "(A ∨ B)," or "(A ⊃ B)." Assume one side or its negation and return to step 2.

5. REFUTE: Construct a refutation box containing any simple wffs (letters or their negation) that aren't blocked off. In the original argument, mark each letter "1" or "0" or "?" depending on whether the box has the letter or its negation or neither. These truth conditions should make the premises all true and conclusion false – thus showing the argument to be invalid.

Let's do another valid one (we'll do invalid later). Here, after deriving a few lines, we get stuck and can't go further. So we need to make another assumption. We could assume the left or right sides (or their denials) of lines 1, 2, or 4.

1 (A ⊃ (B • C))
2 (B ⊃ (A • C))
[∴ ((A ∨ B) ⊃ C)
* 3 asm: ∼((A ∨ B) ⊃ C)
4 ∴ (A ∨ B) {from 3}
5 ∴ ∼C {from 3}

We decide to assume the left side of line 1. Then we derive further lines to get a contradiction (5 and 9). We add double stars, since we have two live assumptions.

```
** 1   (A ⊃ (B • C))
   2   (B ⊃ (A • C))
   [ ∴ ((A ∨ B) ⊃ C)
*  3   asm: ~((A ∨ B) ⊃ C)
   4   ∴ (A ∨ B)    {from 3}
   5   ∴ ~C    {from 3}
   6     asm: A    {break 1}
** 7     ∴ (B • C)    {from 1 and 6}
   8     ∴ B    {from 7}
   9     ∴ C    {from 7}
```

We then block off from assumption 6 down, conclude its opposite in line 10, and (since we now have only one live assumption) erase double stars. As we continue the proof, we ignore blocked-off lines (the original conclusion and 6 to 9). 0167

```
    1   (A ⊃ (B • C))
    2   (B ⊃ (A • C))
    [ ∴ ((A ∨ B) ⊃ C)
*   3   asm: ~((A ∨ B) ⊃ C)
    4   ∴ (A ∨ B)    {from 3}
    5   ∴ ~C    {from 3}
    6   ┌ asm: A    {break 1}
    7   │ ∴ (B • C)    {from 1 and 6}
    8   │ ∴ B    {from 7}
    9   └ ∴ C    {from 7}
   10   ∴ ~A    {from 6; 5 contradicts 9}
```

We then derive further lines and get our second contradiction (lines 10 and 13). We apply RAA again, this time on our original assumption.

```
    1   (A ⊃ (B • C))    Valid
  * 2   (B ⊃ (A • C))
    [ ∴ ((A ∨ B) ⊃ C)
  * 3   ┌ asm: ~((A ∨ B) ⊃ C)
  * 4   │ ∴ (A ∨ B)    {from 3}
    5   │ ∴ ~C    {from 3}
    6   │ ┌ asm: A    {break 1}
    7   │ │ ∴ (B • C)    {from 1 and 6}
    8   │ │ ∴ B    {from 7}
    9   │ └ ∴ C    {from 7}
   10   │ ∴ ~A    {from 6; 5 contradicts 9}
   11   │ ∴ B    {from 4 and 10}
   12   │ ∴ (A • C)    {from 2 and 11}
   13   └ ∴ A    {from 12}
   14   ∴ ((A ∨ B) ⊃ C)    {from 3; 10 contradicts 13}
```

Since every assumption has led to a contradiction, our proof is done.

7.3a Exercise: LogiCola GHV

Prove each of these arguments to be valid (all are valid). 0168

(B ∨ A)
(B ⊃ A)
∴ ~(A ⊃ ~A)

```
* 1  (B ∨ A)   Valid
  2  (B ⊃ A)
     [ ∴ ~(A ⊃ ~A)
* 3  ┌ asm: (A ⊃ ~A)
  4  │ ┌ asm: B   {break 1}
  5  │ │ ∴ A   {from 2 and 4}
  6  │ └ ∴ ~A   {from 3 and 5}
  7  │ ∴ ~B   {from 4; 5 contradicts 6}
  8  │ ∴ A   {from 1 and 7}
  9  └ ∴ ~A   {from 3 and 8}
 10 ∴ ~(A ⊃ ~A)   {from 3; 8 contradicts 9}
```

1. (A ⊃ B)
(A ∨ (A • C))
∴ (A • B)

2. (((A • B) ⊃ C) ⊃ (D ⊃ E))
D
∴ (C ⊃ E)

3. (B ⊃ A)
~(A • C)
(B ∨ C)
∴ (A ≡ B)

4. (A ∨ (D • E))
(A ⊃ (B • C))
∴ (D ∨ C)

5. ((A ⊃ B) ⊃ C)
(C ⊃ (D • E))
∴ (B ⊃ D)

6. $(\sim(A \lor B) \supset (C \supset D))$
$(\sim A \bullet \sim D)$
$\therefore (\sim B \supset \sim C)$

7. $(\sim A \equiv B)$
$\therefore \sim(A \equiv B)$

8. $(A \supset (B \bullet \sim C))$
C
$((D \bullet \sim E) \lor A)$
$\therefore D$

7.3b Exercise: LogiCola GHV

First appraise intuitively. Then translate into logic (using the letters given) and prove to be valid (all are valid).

1. Either the butler fixed the drink and poisoned the deceased, or the butler added poison later and poisoned the deceased.
If the butler poisoned the deceased, then the butler is guilty.
∴ The butler poisoned the deceased and is guilty. [Use F, P, A, and G.]

2. If I'm coming down with a cold and I exercise, then I'll get worse and feel awful.
If I don't exercise, then I'll suffer exercise deprivation and I'll feel awful.
∴ If I'm coming down with a cold, then I'll feel awful. [Use C, E, W, A, and D. This one is easier if you break premise 1 (not premise 2) to make your assumption.]

3. You'll get an A if and only if you either get a hundred on the final exam or else bribe the teacher.
You won't get a hundred on the final exam.
∴ You'll get an A if and only if you bribe the teacher. [Use A, H, and B.] 0169

4. If President Nixon knew about the massive Watergate cover-up, then he lied to the American people on national television and he should resign.
If President Nixon didn't know about the massive Watergate cover-up, then he was incompetently ignorant and he should resign.
∴ Nixon should resign. [K, L, R, I]

5. If you don't compromise your principles, then you won't get campaign money.
If you won't get campaign money, then you won't be elected.
If you compromise your principles, then you'll appeal to more voters.
If you appeal to more voters, then you'll be elected.
∴ You'll be elected if and only if you compromise your principles. [C, M, E, A]

6. Moral judgments express either truth claims or feelings.
If moral judgments express truth claims, then "ought" expresses either a concept from sense experience or an objective concept that isn't from sense experience.
"Ought" doesn't express a concept from sense experience.
"Ought" doesn't express an objective concept that isn't from sense experience.
∴ Moral judgments express feelings and not truth claims. [T, F, S, O]

7. If Michigan either won or tied, then Michigan is going to the Rose Bowl and Gensler is happy.
∴ If Gensler isn't happy, then Michigan didn't tie. [W, T, R, H]

8. There are moral obligations.
If there are moral obligations and moral obligations are explainable, then either there's an explanation besides God's existence or else God's existence would explain moral obligations.
God's existence wouldn't explain moral obligation.
∴ Either moral obligations aren't explainable, or else there's an explanation besides God's existence. [M, E, B, G]

9. If determinism is true and Dr Freudlov correctly predicts (using deterministic laws) what I'll do, then if she tells me her prediction I'll do something else.
If Dr Freudlov tells me her prediction and yet I'll do something else, then Dr Freudlov doesn't correctly predict (using deterministic laws) what I'll do.
∴ If determinism is true, then Dr Freudlov doesn't correctly predict (using deterministic laws) what I'll do or else she won't tell me her prediction. [D, P, T, E]

10. If you make this demand on your son [that he leave Suzy or else not have his graduate schooling financed] and he leaves Suzy, then he'll regret being forced to leave her and he'll always resent you.
If you make this demand on your son and he doesn't leave Suzy, then he'll regret not going to graduate school and he'll always resent you.
∴ If you make this demand on your son, then he'll always resent you. [Use D, L, F, A, and G; this one is difficult.] 0170

7.4 Harder refutations

With multiple-assumption *invalid* arguments, we keep making assumptions until we get our refutation. Here's an example:

> If the butler was at the party, he fixed the drinks and poisoned the deceased.
> If the butler wasn't at the party, he was at a neighbor's house.

∴ The butler poisoned the deceased.

1 $(A^0 \supset (F^? \bullet P^0)) = 1$ **Invalid**
** 2 $(\sim A^0 \supset N^1) = 1$
[∴ $P^0 = 0$
3 asm: ~P
4 asm: ~A {break 1}
5 ∴ N {from 2 and 4}

N, ~A, ~P

We derive all we can and make additional assumptions when needed. We reach a refutation in which the butler was at a neighbor's house, wasn't at the party, and didn't poison the deceased. This makes the premises all true and conclusion false.

Follow the five-step proof strategy of the previous section until you get a proof or a refutation. If every assumption leads to a contradiction, then you get a proof. But when do you know that the argument is invalid? When do you stop making further assumptions and instead construct a refutation box? *Stop and refute when you can't derive anything further (using S- or I-rules or RAA) and every complex wff is starred or blocked off or broken.* (A complex wff is "broken" if we have one side or its negation but not what we need to conclude anything new.)

This invalid argument requires three assumptions:

1 $(A^0 \supset B^?) = 1$ **Invalid**
2 $(C^0 \supset D^?) = 1$
3 $(F^0 \supset (C^0 \bullet D^?)) = 1$
[∴ $(E^1 \supset C^0) = 0$
* 4 asm: ~(E ⊃ C)
5 ∴ E {from 4}
6 ∴ ~C {from 4}
7 asm: ~A {break 1}
8 asm: ~F {break 3}

E, ~A, ~C, ~F

Here we can derive nothing further and all complex wffs are either starred (line 4), blocked off (original conclusion), or broken (lines 1–3). Our refutation, even without values for "B" or "D," makes the premises all true and conclusion false.

Our proof strategy, if applied correctly, will always give a proof or refutation. How these go may depend on which lines we do first and what we decide to assume; proofs and refutations may differ but still be correct. 0171

7.4a Exercise: LogiCola GHI

Prove each of these arguments to be invalid (all are invalid).

(A ∨ ~(B ⊃ C))
(D ⊃ (A ⊃ B))
∴ (C ⊃ ~(D ∨ A))

1 $(A^{1} \vee \sim(B^{?} \supset C^{1})) = 1$ **Invalid**
2 $(D^{0} \supset (A^{1} \supset B^{?})) = 1$
[$\therefore (C^{1} \supset \sim(D^{0} \vee A^{1})) = 0$
* 3 asm: ~(C ⊃ ~(D ∨ A))
4 ∴ C {from 3}
5 ∴ (D ∨ A) {from 3}
6 asm: A {break 1}
7 asm: ~D {break 2}

A, C, ~D

1. ~(A • B)
∴ (~A • ~B)

2. (A ⊃ ~B)
∴ ~(A ⊃ B)

3. (A ⊃ B)
(C ⊃ (~D • E))
∴ (D ∨ F)

4. ~(A • B)
∴ ~(A ≡ B)

5. (A ⊃ (B • C))
((D ⊃ E) ⊃ A)
∴ (E ∨ C)

6. (~A ∨ ~B)
∴ ~(A ∨ B)

7. ((A • B) ⊃ ~(C • D))
C
(E ⊃ B)
∴ ~E

8. $(A \supset (B \supset C))$
$(B \vee \sim(C \supset D))$
$\therefore (D \supset \sim(A \vee B))$

7.4b Exercise: LogiCola G (HC & MC)

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation).

1. If the maid prepared the drink, then the butler didn't prepare it.
The maid didn't prepare the drink.
If the butler prepared the drink, then he poisoned the drink and is guilty.
∴ The butler is guilty.   [Use M, B, P, and G.]

2. If you tell your teacher that you like logic, then your teacher will think that you're insincere and you'll be in trouble.
If you don't tell your teacher that you like logic, then your teacher will think that you dislike logic and you'll be in trouble.
∴ You'll be in trouble.   [Use L, I, T, and D.]

3. If we don't get reinforcements, then the enemy will overwhelm us and we won't survive.
∴ If we do get reinforcements, then we'll conquer the enemy and we'll survive. [Use R, O, S, and C.] 0172

4. If Socrates didn't approve of the laws of Athens, then he would have left Athens or would have tried to change the laws.
If Socrates didn't leave Athens and didn't try to change the laws, then he agreed to obey the laws.
Socrates didn't leave Athens.
∴ If Socrates didn't try to change the laws, then he approved of the laws and agreed to obey them.   [Use A, L, C, and O; from Plato's *Crito*, which argued that Socrates shouldn't disobey the law by escaping from jail.]

5. If I hike the Appalachian Trail and go during late spring, then I'll get maximum daylight and maximum mosquitoes.
If I get maximum mosquitoes, then I won't be comfortable.
If I go right after school, then I'll go during late spring.
∴ If I hike the Appalachian Trail and don't go right after school, then I'll be comfortable.   [A, L, D, M, C, S]

6. [Logical positivism says "*Every genuine truth claim is either experimentally testable or true by definition*." This view, while once popular, is self-refuting and hence not very popular today.]
If LP (logical positivism) is true and is a genuine truth claim, then it's either experimentally testable or true by definition.
LP isn't experimentally testable.
LP isn't true by definition.
If LP isn't a genuine truth claim, then it's not true.
∴ LP isn't true. [T, G, E, D]

7. If you give a test, then students either do well or do poorly.
If students do well, then you think you made the test too easy and you're frustrated.
If students do poorly, then you think they didn't learn any logic and you're frustrated.
∴ If you give a test, then you're frustrated. [Use T, W, P, E, F, and L; from a class who tried to talk me out of giving a test.]

8. If the world contains moral goodness, then the world contains free creatures and the free creatures sometimes do wrong.
If the free creatures sometimes do wrong, then the world is imperfect and the creator is imperfect.
∴ If the world doesn't contain moral goodness, then the creator is imperfect. [M, F, S, W, C]

9. We'll find your action's cause, if and only if your action has a cause and we look hard enough.
If all events have causes, then your action has a cause.
All events have causes.
∴ We'll find your action's cause, if and only if we look hard enough. [F, H, L, A]
0173

10. Herman sees that the piece of chalk is white.
The piece of chalk is the smallest thing on the desk.
Herman doesn't see that the smallest thing on the desk is white. (He can't see the whole desk and so can't tell that the piece of chalk is the smallest thing on it.)
If Herman sees a material thing, then if he sees that the piece of chalk is white and the piece of chalk is the smallest thing on the desk, then he sees that the smallest thing on the desk is white.
If Herman doesn't see a material thing, then he sees a sense datum.
∴ Herman doesn't see a material thing, but he does see a sense datum. [Use H, P, H´, M, and S. This argument attacks direct realism: that we directly perceive material objects and not just sensations.]

11. If the final capacitor in the transmitter is arcing, then the SWR (standing wave ratio) is too high and the efficiency is lowered.
If you hear a cracking sound, then the final capacitor in the transmitter is arcing.
∴ If you don't hear a cracking sound, then the SWR isn't too high. [A, H, L, C]

12. If we can know that God exists, then we can know God by experience or we can know God by logical inference from experience.
If we can't know God empirically, then we can't know God by experience and we can't know God by logical inference from experience.
If we can know God empirically, then "God exists" is a scientific hypothesis and is empirically falsifiable.
"God exists" isn't empirically falsifiable.
∴ We can't know that God exists. [K, E, L, M, S, F]

13. If I perceive, then my perception is either delusive or veridical.
If my perception is delusive, then I don't directly perceive a material object.
If my perception is veridical and I directly perceive a material object, then my experience in veridical perception would always differ qualitatively from my experience in delusive perception.
My experience in veridical perception doesn't always differ qualitatively from my experience in delusive perception.
If I perceive and I don't directly perceive a material object, then I directly perceive a sensation.
∴ If I perceive, then I directly perceive a sensation and I don't directly perceive a material object. [Use P, D, V, M, Q, and S. This argument from illusion attacks direct realism: that we directly perceive material objects and not just sensations.]

14. If you're romantic and you're Italian, then Juliet will fall in love with you and will want to marry you.
If you're Italian, then you're romantic.
∴ If you're Italian, then Juliet will want to marry you. [R, I, F, M]

15. If emotions can rest on factual errors and factual errors can be criticized, then we can criticize emotions.
If we can criticize emotions and moral judgments are based on emotions, then beliefs about morality can be criticized and morality isn't entirely non-rational.
∴ If morality is entirely non-rational, then emotions can't rest on factual errors. [E, F, W, M, B, N] 0174

7.5 Copi proofs

There are many proof methods for propositional logic. *Copi proofs* are based on an early and still popular method. Copi proofs use a somewhat standard set of inference and replacement rules.[1] These eight *inference rules*, like our S- and I-rules, let us infer whole lines from previous whole lines (here each capital letter may be uniformly replaced by any wff):

AD Addition P —————— (P ∨ Q)
CJ Conjunction P Q ——————— (P • Q)
DI Dilemma ((P ⊃ Q) • (R ⊃ S)) (P ∨ R) ———————————————— (Q ∨ S)
DS Disjunctive Syllogism (P ∨ Q) ~P ——————— Q
HS Hypothetical Syllogism (P ⊃ Q) (Q ⊃ R) ——————— (P ⊃ R)
MP Modus Ponens (P ⊃ Q) P ——————— Q

[1] This proof method goes back to Irving Copi's *Introduction to Logic* (New York: Macmillan, 1953) and has appeared with variations in many books. Copi's original list (his p. 259) also had Destructive Dilemma ("((P ⊃ Q) • (R ⊃ S)), (~Q ∨ ~S) ∴ (~P ∨ ~R)") but omitted Repetition's second part. Absorption ("(P ⊃ Q) ∴ (P ⊃ (P • Q))" and "(P ⊃ (P • Q)) ∴ (P ⊃ Q)") was sometimes added later. I simplified some names and gave each rule a two-letter abbreviation.

MT Modus Tollens
(P ⊃ Q) ~Q ------- ~P
SP Simplification
(P • Q) ------- P

To explore how these work, we'll compare them to our S- and I-rules.
Our first three S-rules are AND, NOR, and NIF:

AND	NOR	NIF
(P • Q) ------ P, Q	~(P ∨ Q) ------- ~P, ~Q	~(P ⊃ Q) ------- P, ~Q

Copi can derive the AND rule. We can get the left side of "(P • Q)" by using SP (Simplification) directly on line 1, as indicated by the "{SP 1}" justification:

1 (P • Q)
2 P {SP 1}

To get the right side, we first switch sides to get "(Q • P)," using replacement rule CM (Commutation), which we'll present later. Then we use SP:

1 (P • Q)
2 (Q • P) {CM 1)
3 Q (SP 2} 0175

Copi can also derive the NOR conclusions. To get the left side, we first apply the DM (De Morgan) replacement rule, which we'll present later, to go from "~(P ∨ Q)" to "(~P • ~Q)." Then we use SP to derive "~P":

1 ~(P ∨ Q)
2 (~P • ~Q) {DM 1}
3 ~P {SP 2}

To get the right side, we use similar reasoning, but we have to again switch sides using CM:

1 ~(P ∨ Q)
2 (~P • ~Q) {DM 1}
3 (~Q • ~P) {CM 2}
4 ~Q {SP 3}

Deriving NIF is more involved. To get the left side, we first reshape the IF-THEN into an OR (using Implication replacement rule IM) and then an AND (using De Morgan replacement rule DM). We apply SP to get "~~P," and then the double-negation replacement rule DN to get "P":

1 ~(P ⊃ Q)
2 ~(~P ∨ Q) {IM 1}
3 (~~P • ~Q) {DM 2}
4 ~~P {SP 3}
5 P {DN 4}

Getting the right side is similar, but we switch sides using CM before using SP to get "~Q":

1 ~(P ⊃ Q)
2 ~(~P ∨ Q) {IM 1}
3 (~~P • ~Q) {DM 2}
4 (~Q • ~~P) {CM 3}
5 ~Q {SP 4}

These examples show how difficult the Copi method can be to use. But that's the challenge – it makes us think hard about how to derive a conclusion, and maybe think out various possible approaches first; some teachers like Copi proofs for exactly this reason.

Replacement rules are important in Copi proofs. These ten *replacement rules* let you switch *one* occurrence of identical formulas anywhere in a wff:

AS Association $(P \lor (Q \lor R)) = ((P \lor Q) \lor R)$ $(P \bullet (Q \bullet R)) = ((P \bullet Q) \bullet R)$
CM Commutation $(P \lor Q) = (Q \lor P)$ $(P \bullet Q) = (Q \bullet P)$
DB Distribution $(P \bullet (Q \lor R)) = ((P \bullet Q) \lor (P \bullet R))$ $(P \lor (Q \bullet R)) = ((P \lor Q) \bullet (P \lor R))$
DM De Morgan $\sim(P \bullet Q) = (\sim P \lor \sim Q))$ $\sim(P \lor Q) = (\sim P \bullet \sim Q)$
DN Double Negation $P = \sim\sim P$
EQ Equivalence $(P \equiv Q) = ((P \supset Q) \bullet (Q \supset P))$ $(P \equiv Q) = ((P \bullet Q) \lor (\sim P \bullet \sim Q))$
EX Exportation $((P \bullet Q) \supset R) = (P \supset (Q \supset R))$
IM Implication $(P \supset Q) = (\sim P \lor Q)$ 0176
RP Repetition $P = (P \lor P)$ $P = (P \bullet P)$
TR Transposition $(P \supset Q) = (\sim Q \supset \sim P)$

These reshape formulas to fit the inference rules.

Let's consider how the Copi method can mirror our I-rules:

NOT-BOTH		OR		IF-THEN	
~(P • Q) P ------ ~Q	~(P • Q) Q ------ ~P	(P ∨ Q) ~P ------ Q	(P ∨ Q) ~Q ------ P	(P ⊃ Q) P ------ Q	(P ⊃ Q) ~Q ------ ~P

OR conclusions are easy to derive using DS (Disjunctive Syllogism); for the second version, we also switch sides using CM:

1 (P ∨ Q)
2 ~P
3 Q {DS 1+2}

1 (P ∨ Q)
2 ~Q
3 (Q ∨ P) {CM 1}
4 P {DS 2+3}

NOT-BOTH uses DS with DM (De Morgan):

1 ~(P • Q)
2 P
3 (~P ∨ ~Q) {DM 1}
4 ~~P {DN 2}
5 ~Q {DS 3+4}

1 ~(P • Q)
2 Q
3 (~P ∨ ~Q) {DM 1}
4 (~Q ∨ ~P) {CM 3}
5 ~~Q {DN 2}
6 P {DS 4+5}

Since Copi rules take "not" very strictly, we can't on the left go directly from "P" and "(~P ∨ ~Q)" to get "~Q"; instead, we have to double negate "P" to get "~~P," which is like the first part but starts with an additional squiggle. MP (Modus Ponens) and MT (Modus Tollens) parallel our IF-THEN forms.

Here's a Copi proof for the butler example in §7.1:

Conclusion: B

1 T
2 (T ⊃ (B ∨ M))
3 (M ⊃ H)
4 ~H

5 (B ∨ M) {MP 1+2}
6 ~M {MT 3+4}
7 (M ∨ B) {CM 5}
8 B {DS 6+7} 0177

If we try to prove the invalid butler example in §7.2, we won't derive the conclusion; but this may just be due to our lack of ingenuity. The Copi method won't show invalid arguments to be invalid; so it's normally used only on arguments already known to be valid, which limits the method's usefulness.

Conclusion: B

1 T
2 (T ⊃ (B ∨ M))
3 (M ⊃ H)
4 (B ∨ M) {MP 1+2}
? ? ?

Here's a Copi proof for the multiple-assumption butler example in §7.3:

Conclusion: P

1 (A ⊃ (F • P))
2 (~A ⊃ (S • R))
3 ~R
4 (~R ∨ ~S) {AD 3}
5 ~(R • S) {DM 4}
6 ~(S • R) {CM 5}
7 ~~A {MT 2+6}
8 A {DN 7}
9 (F • P) {MP 1+8}
10 (P • F) {CM 9}
11 P {SP 10}

Lines 4 to 6 use a common strategy: think of what wff we need – here we need "~(S • R)" to use with line 2 and MT – and how to get it from what we have – here "~R" can provide "(~R ∨ ~S)," which we reshape into "~(S • R)."

So far, we've used Copi *direct proofs*, where the conclusion is derived from the premises without making any assumptions. Copi also provides for *conditional proofs* and *indirect proofs [reductio ad absurdum]* (using CP and RA):

CP Conditional Proof

If you assume P and later derive Q, then you can star all the lines from P to Q [showing that you aren't to use them to derive further steps] and then derive (P ⊃ Q).

RA Reductio ad Absurdum

If you assume P and later derive (Q • ~Q), then you can star all the lines from P to (Q • ~Q) [showing that you aren't to use them to derive further steps] and then derive ~P.

The proof isn't done until all assumptions are starred. Here are examples (add "*" when applying RA or CP; ignore starred lines in deriving further steps):

Conclusion: ((A • B) ⊃ A)

1 (A • B) {Assume} *
2 A {SP 1} *
3 ((A • B) ⊃ A) {CP 1+2}

Conclusion: (A ∨ ~A)

1 ~(A ∨ ~A) {Assume} *
2 (~A • ~~A) {DM 1} *
3 ~~(A ∨ ~A) {RA 1+2}
4 (A ∨ ~A) {DN 3} 0178

CP and RA are useful in proving logical truths from zero premises. CP is convenient for proving conditional conclusions. And if you're really confused on how to do a problem, I suggest that you start by assuming the conclusion's opposite; try to derive a contradiction and then apply RA.

Comparing our method to Copi's, all the same arguments are provable. Our method is easier to learn (with a smaller and more systematic set of rules), easier to use (with a proof procedure that doesn't require guesswork or intuition), and more powerful (since it can refute invalid arguments). But you might want to learn the Copi method too; Copi proofs are good mental exercise and can be fun (especially on LogiCola) – and Copi rules are sometimes assumed in philosophical discussions.

On LogiCola, you do Copi proofs by picking "Copi Proofs" and the level of difficulty (Easier / Harder / Mixed); you get the same randomly generated problems (but only valid ones) as with our usual proofs. You repeatedly type the next wff, click (or touch) the inference rule, and then click (or touch) the previous wffs from which your step follows. There are no arrows to get the next line or finish the problem; but you can quit the problem (which costs you points) or paste your own problems (or ones from your teacher). You can also copy previous lines or the conclusion into the answer space, so you can modify them to give your next line. While Copi proofs are difficult, you'll soon get the hang of it.

7.5a and 7.5b Exercise: LogiCola GEO

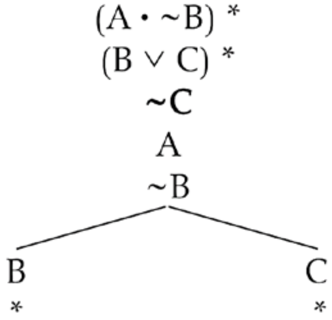
Do Copi proofs for problems in §§7.1a and 7.1b (all are valid). These are easier problems.

7.5c and 7.5d Exercise: LogiCola GHO and GMO

Do Copi proofs for problems in §§7.3a and 7.3b (all are valid). These are harder problems.

7.6 Truth trees

Also common are *truth trees*, which break formulas into the cases that make them true. Here's a truth tree for "(A • ~B), (B ∨ C) ∴ C" – which comes out as valid, because every branch closes:



0179 First write the premises and the conclusion's contradictory. Then break the complex formulas into the cases that make them true, to see if there's some way to get premises all true and conclusion false. Simplify "(A • ~B)" into "A" and "~B" and then star it (as broken). Branch "(B ∨ C)" into the two cases that make it true and then star it (as broken); so one branch has "B" and another has "C." Both branches are self-contradictory, since the first has "B" on the branch and "~B" on the trunk – and the second has "C" on the branch and "~C" on the trunk; *close* both branches by adding "*" to the bottom. The argument is VALID, since having premises all true and conclusion false is impossible.

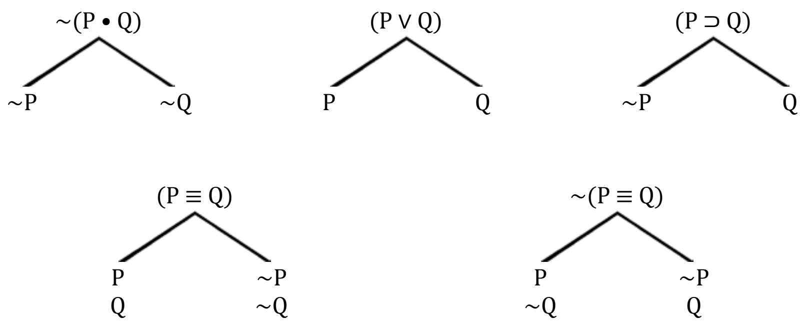
Truth trees use simplifying and branching rules (both apply only to whole lines). These *simplifying rules* simplify a formula into smaller parts:

$$\frac{(P \bullet Q)}{P,\ Q} \qquad\qquad \frac{\sim(P \vee Q)}{\sim P,\ \sim Q}$$

$$\frac{\sim(P \supset Q)}{P,\ \sim Q} \qquad\qquad \frac{\sim\sim P}{P}$$

When you use these, put a star after the original formula to show that it's *broken* (this means that its truth is assured by the truth of some smaller parts below). Use simplifying rules before branching rules (this is more efficient).

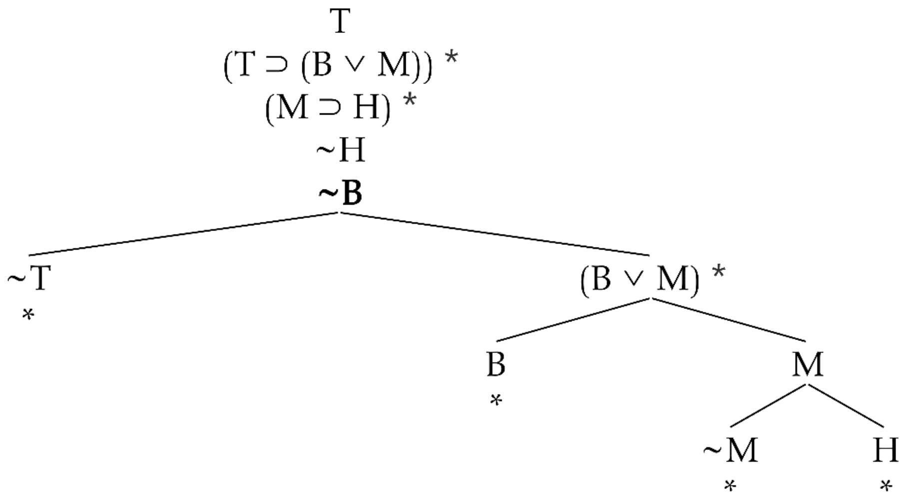
These *branching rules* branch a formula into the two sub-cases that would make it true (so "~(P • Q)" is true just if "~P" is true or "~Q" is true):



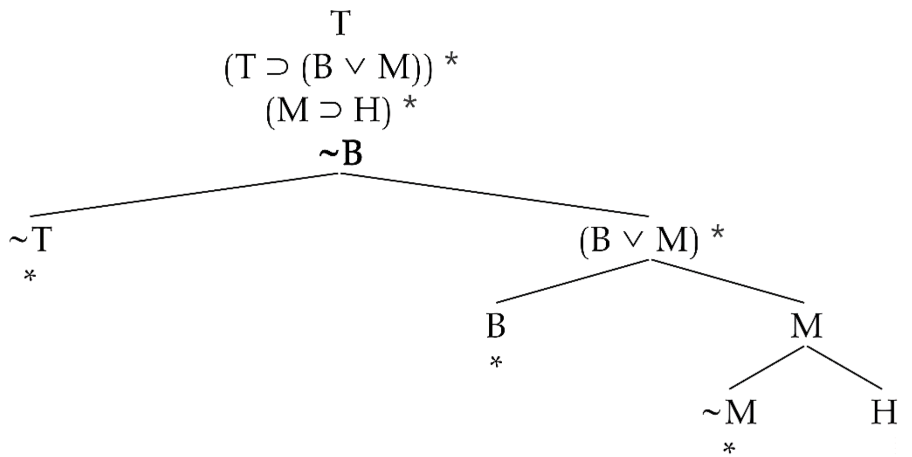
When you use these, put a star after the original formula to show that it's broken; add the sub-branches to the bottom of any bottom branch further down that isn't yet marked as closed (self-contradictory).

To test an argument, write the premises and the conclusion's contradictory. Keep applying the simplifying and branching rules to complex unstarred formulas. Close a branch when it has contradictory formulas; a closed branch is a failed attempt to make the premises all true and conclusion false. If all branches close, then the argument is valid. If some branch doesn't close and yet all its complex wffs are starred (broken), then there's a possible way to get premises all true and conclusion false, and so the argument is invalid. 0180

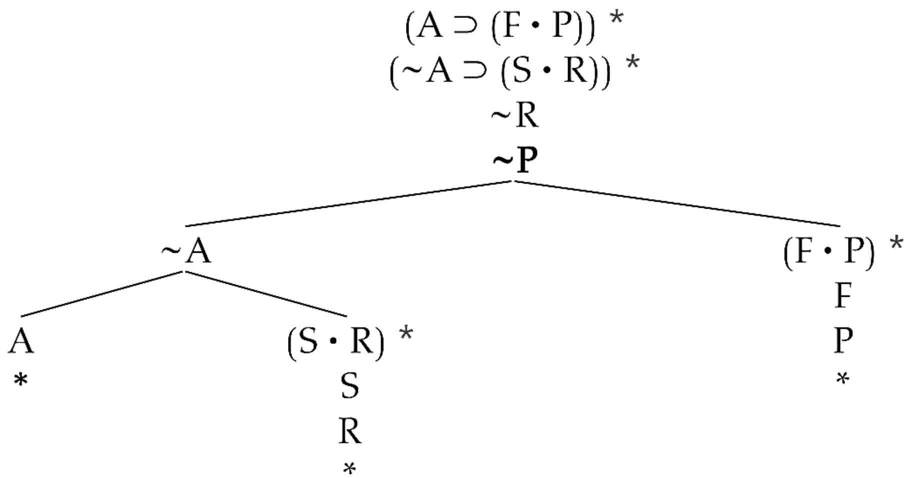
To show further how this works, I'll do truth trees for the butler examples in §§7.1–7.4. The argument in §7.1 comes out valid, since every branch closes:



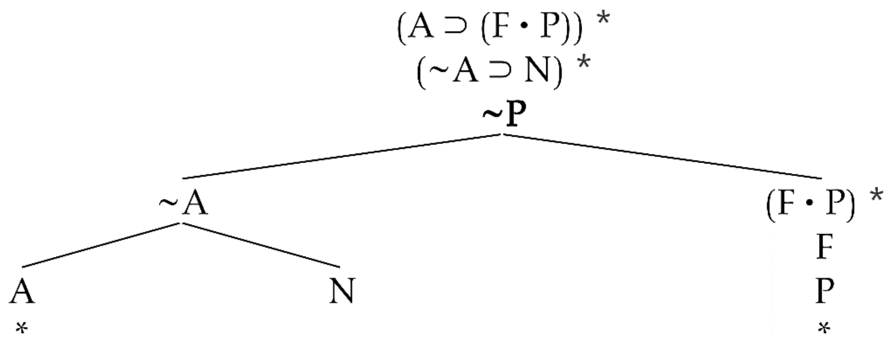
Here we branch line 2 into "~T" and "(B ∨ M)," branch the latter into "B" and "M," and branch "(M ⊃ H)" into "~M" and "H"; every branch closes, and so it's valid. The butler argument in §7.2 is invalid, since without "~H" the branch ending in "H" doesn't close (the refutation has: T, M, H, and ~B):



The butler argument in §7.3 comes out valid, since every branch closes:



0181 The butler argument in §7.4 comes out invalid, since the branch ending in “N” doesn’t close (the refutation has: N, ~A, and ~P):



As compared with Copi proofs, truth trees are easier to do (since they use an easily learned strategy) and can test for validity or invalidity. But truth trees don’t mirror ordinary reasoning as well; they give a mechanical way to test validity instead of a way to help develop reasoning skills. Our proof method tries to combine the strengths of both methods. Like truth trees, our proofs have an easily learned strategy, keep breaking formulas into simpler parts, and can test for validity or invalidity. But like Copi proofs, our proofs give a linear derivation of formulas that somewhat mirrors ordinary reasoning. Our proofs use similar simplification rules as truth trees, but replace branching with inference rules and assumptions.

On LogiCola, you do truth trees by picking “Treez” and the level of difficulty (Easier / Harder / Mixed); you get the same randomly generated problems (valid and invalid) as with regular proofs. You do this exercise entirely by clicking or touching (no typing); follow the directions at the bottom. The "program closes branches" option automatically closes self-contradictory branches, while "you close branches" has you close these yourself (but without losing points for errors). The "automatic double-negation" option simplifies double negations automatically (so "∼(A ∨ ∼B)" simplifies into "∼A" and "B" – instead of "∼A" and "∼∼B"). You can click (or touch) arrows at the top to give you the next line or finish the problem (but without getting credit for the problem); and you can use these arrows to step through sample proofs before doing them on your own.

7.6a Exercise: LogiCola GEZ

Do truth trees for problems in §§7.1a, 7.1b, 7.2a, and 7.2b. These are easier problems.

7.6b Exercise: LogiCola GHZ and GMZ

Do truth trees for problems in §§7.3a, 7.3b, 7.4a, and 7.4b. These are harder problems.

8 Basic Quantificational Logic

Quantificational logic, which builds on propositional logic, studies arguments whose validity depends on notions like "all," "no," and "some." This system is stronger than syllogistic logic (Chapter 2), since it can express complex ideas like "If some are A, then all that are B or C are then D but not E." This chapter covers the basics and the next adds relations and identity.

8.1 Easier translations

To help us evaluate arguments, we'll construct a quantificational language. This will include propositional logic's vocabulary, wffs, inference rules, and proofs. It adds two new vocabulary items: small letters and "∃." Here are sample wffs:

Ir	=	Romeo is Italian.
Ix	=	x is Italian.
(x)Ix	=	For all x, x is Italian (all are Italian).
(∃x)Ix	=	For some x, x is Italian (some are Italian).

Learn to express "All are Italian" as "For all x, x is Italian." This uses *Loglish*, a mix of logic and English. Loglish helps us to translate from English to logic.

"Romeo is Italian" is "Ir"; the capital letter goes first. "I" is for the general category "Italian" and "r" is for the specific individual "Romeo":

Use capital letters for **general terms**, which *describe* or put in a *category*:

I = an Italian
C = charming
F = drives a Ford

Use capitals for "a so and so," adjectives, and verbs.

Use small letters for **singular terms**, which pick out a *specific* person or thing:

i = the richest Italian
t = this child
r = Romeo

Use small letters for "the so and so," "this so and so," and proper names.

Letters here have various uses. Capitals can represent statements, general
0183 terms, or relations (which we take in the next chapter):

A capital letter alone (not followed by small letters) represents a *statement*.

S = *It's snowing*.

A capital letter followed by a single small letter represents a *general term*.

Ir = Romeo is *Italian*.

A capital letter followed by two or more small letters represents a *relation*.

Lrj = Romeo *loves* Juliet.

Small letters can be constants or variables:

A small letter from "a" to "w" is a **constant** (it refers to a specific person or thing).

Ir = *Romeo* is Italian.

A small letter from "x" to "z" is a **variable** (its reference isn't directly specified).

Ix = *x* is Italian.

"Ix" ("x is Italian") is incomplete, and so not true or false, since we haven't said whom we're talking about. Quantifiers can complete the claim. A **quantifier** is a sequence of the form "(x)" or "(∃x)" – where any variable may replace "x":

“(x)” is a **universal quantifier**. It says that the next formula is true for *all* values of x.

(x)Ix = For all x, x is Italian (all are Italian).

“(∃x)” is an **existential quantifier**. It says that the next formula is true for *at least one* value of x.

(∃x)Ix = For some x, x is Italian (some are Italian).

Quantifiers express “all” and “some” by saying in how many cases the following formula is true.

As before, grammatical formulas are *wffs* (*well-formed formulas*). Wffs now are strings we can construct using the propositional rules plus two new rules:

1. The result of writing a capital letter and then a small letter is a wff.
2. The result of writing a quantifier and then a wff is a wff.

These rules let us build wffs that we’ve already mentioned: “Ir,” “Ix,” “(x)Ix,” and “(∃x)Ix.” Don’t use additional parentheses; these forms are incorrect: “(Ir),” “(Ix),” “(x)(Ix),” “(∃x)(Ix),” “((x)Ix),” “((∃x)Ix).” Use a pair of parentheses for each quantifier and each instance of “•,” “∨,” “⊃,” and “≡”; use no other parentheses. Here are some further wffs: 0184

∼(x)Ix = Not all are Italian
It’s false that, for all x, x is Italian

∼(∃x)Ix = No one is Italian
It’s false that, for some x, x is Italian

(Ix ⊃ Lx) = If x is Italian then x is a lover

(Ix • Lx) = x is Italian and x is a lover

Translating from English to wffs can be difficult. We’ll begin with sentences that translate into wffs starting with a quantifier, or with “∼” and then a quantifier. This rule tells where to put what quantifier:

If the English begins with "**all**" or "**every**," then begin the wff with "**(x)**."
If the English begins with "**not all**" or "**not every**," then begin the wff with "**∼(x)**."
If the English begins with "**some**," then begin the wff with "**(∃x)**."
If the English begins with "**no**," then begin the wff with "**∼(∃x)**."

All are Italian = (x)Ix
Not all are Italian = ∼(x)Ix
Some are Italian = (∃x)Ix
No one is Italian = ∼(∃x)Ix

Here are harder examples:

All are rich or Italian
= (x)(Rx ∨ Ix)

Not everyone is non-Italian
= ∼(x)∼Ix

Some aren't rich
= (∃x)∼Rx

No one is rich and non-Italian
= ∼(∃x)(Rx • ∼Ix)

When the English begins with "all," "not all," "some," or "no," put the quantifier *outside* all parentheses. So "All are rich or Italian" is "(x)(Rx ∨ Ix)." Don't translate it as "((x)Rx ∨ Ix)," which means "Either everyone is rich, or x is Italian."

If the English sentence uses a word like "or," "and," or "if-then," then use the corresponding logical symbol. Otherwise, follow these rules:

With "all ... is ...," use "⊃" for the *middle* connective.
Otherwise use "•" for the connective.

All Italians are lovers
= (x)(Ix ⊃ Lx)
For all x, *if* x is Italian *then* x is a lover 0185

Some Italians are lovers
= (∃x)(Ix • Lx)

For some x, x is Italian *and* x is a lover

No Italians are lovers
= ∼(∃x)(Ix • Lx)
It's false that, for some x, x is Italian *and* x is a lover

With "All Italians ...," think "For all x, *if* x is Italian *then* ...." With "Some Italians ...," think "For some x, x is Italian *and* ...." This example is harder:

All rich Italians are lovers
= (x)((Rx • Ix) ⊃ Lx)
For all x, *if* x is rich *and* Italian, *then* x is a lover

Here use "⊃" as the *middle* connective ("*If* rich Italian, *then* lover") and "•" in the *other* place ("If rich *and* Italian, then lover"). Here are further examples:

Not all Italians are lovers
= ∼(x)(Ix ⊃ Lx)
It's false that, for all x, *if* x is Italian *then* x is a lover

All are rich Italians
= (x)(Rx • Ix)
For all x, x is rich *and* Italian

Sometimes we must rephrase to make "is" (or "are") the main verb:

All dogs hate cats
= All dogs are cat-haters
= (x)(Dx ⊃ Hx)
For all x, if x is a dog then x is a cat-hater

In case of doubt, say the formula in Loglish and see if it means the same as the English sentence. Our translation rules are rough and don't always work.

The **universe of discourse** is the set of entities that words like "all" "some," and "no" range over in a given context. Restricting the universe of discourse to one kind of entity (such as persons or statements) can simplify how we translate some arguments. We'll often restrict the universe of discourse to persons. We did this implicitly when we translated "All are Italian" as "(x)Ix" instead of "(x)(Px ⊃ Ix)" ("All persons are Italian").

Since quantificational translations are so difficult, LogiCola gives you the option to start off by having Loglish hints for these problems.

8.1a Exercise: LogiCola H (EM & ET)

Translate these English sentences into wffs.

Not all logicians run.
∼(x)(Lx ⊃ Rx)

1. x isn't a cat.
2. Something is a cat.
3. Something isn't a cat. 0186
4. It's false that there is something that isn't a cat.
5. Everything is a cat.
6. If x is a dog, then x is an animal.
7. All dogs are animals.
8. No one is evil.
9. Some logicians are evil.
10. No logician is evil.
11. All black cats are unlucky.
12. Some dogs are large and hungry.
13. Not all hungry dogs bark.
14. Some animals aren't barking dogs.
15. Some animals are non-barking dogs.
16. All dogs who bark are frightening.
17. Not all non-dogs are cats.
18. Some cats who aren't black are unlucky.
19. Some cats don't purr.
20. Not every cat purrs.
21. Not all animals are dogs or cats.
22. All who are either dogs or cats are animals.
23. All who are both dogs and cats are animals.
24. All dogs and cats are animals.
25. Everyone is a crazy logician.

8.2 Easier proofs

We need quantifier inference rules. The reverse-squiggle rules hold regardless of what variable replaces "x" and what pair of contradictory wffs replaces "Fx" / "~Fx"; here "→" means that we can infer whole lines from left to right:

Reverse squiggle RS		
~(x)Fx	→	(∃x)~Fx
~(∃x)Fx	→	(x)~Fx

"Not everyone is funny" entails "Someone isn't funny." And "It's false that someone is funny" ("No one is funny") entails "Everyone is non-funny." Our rules cover reversing squiggles on longer formulas, if the whole formula begins with "~" and then a quantifier. Here are two examples:

~(∃x)~Gx
∴ (x)~~Gx

~(x)(Lx • ~Mx)
∴ (∃x)~(Lx • ~Mx)

In the first example, we also could conclude "(x)Gx" (dropping "~~"). This next example is illegal in our system, since it fits poorly into our proof strategy, even though it's logically correct:

Don't do this:

(Ir ⊃ ~(x)Gx)
∴ (Ir ⊃ (∃x)~Gx)

0187 Reverse squiggles whenever you have a wff that begins with "~" and then a quantifier; this moves a quantifier to the beginning of the formula, so we can drop it later.

Drop quantifiers using the next two rules (which hold regardless of what variable replaces "x" and what wffs replace "Fx" / "Fa" – provided that the two wffs are identical except that wherever the variable occurs freely[1] in the

[1] An instance of a variable occurs *freely* if it's not part of a wff that begins with a quantifier using that variable; just the first instance of "x" in "(Fx • (x)Gx)" occurs freely. So we'd go from "(∃x)(Fx • (x)Gx)" to "(Fa • (x)Gx)."

former the same constant occurs in the latter). Here's the drop-existential rule:

Drop existential DE
(∃x)Fx → Fa,
use a *new constant*

Suppose *someone* robbed the bank; we can give this person an *arbitrary name* that we make up (like "Al"). Likewise, when we drop an existential, we'll name this "someone" with a *new constant* – one that hasn't yet occurred in earlier lines of the proof.[1] In proofs, we'll use the next unused constant in alphabetical order – starting with "a," then "b," and so on. So if we drop two existentials, then we introduce two new constants:

(∃x)Mx
<u>(∃x)Fx</u>
∴ Ma
∴ Fb

Someone is male, someone is female; let's call the male "a" and the female "b." It's OK to use "a" in the first inference, since it occurs in no earlier line. But the second inference must use "b," since "a" has now already occurred.

We can drop existentials from complicated formulas if the quantifier begins the wff and we replace the variable with the same *new constant* throughout. So this first inference is fine:

<u>(∃x)(Fx • Gx)</u>
∴ (Fa • Ga)

This next example is wrong (because it drops the quantifier using two different constants):

<u>(∃x)(Fx • Gx)</u>
∴ (Fa • Gb)

[1] If more than one person robbed the bank; then our name (or constant) will refer to a random *one* of the robbers. Using a new name is consistent with the robber being mentioned earlier in the argument; different names (like "Al" and "Smith") might refer to the same individual. DE should be used only when there's at least one not-blocked-off assumption; otherwise, the symbolic version of "Someone is a thief, so Gensler is a thief" would be a two-line proof.

This next example is also wrong (since the formula doesn't *begin* with a quantifier – instead it begins with a left-hand parenthesis):

((∃x)Fx ⊃ P)
∴ (Fa ⊃ P)

Drop only initial quantifiers.

Here's the drop-universal rule:

Drop universal DU
(x)Fx → Fa, use any constant

0188 If *everyone* is funny, then Al is funny, Bob is funny, and so on. From "(x)Fx" we can derive "Fa," "Fb," and so on – using any constant. However, it's bad strategy to use a new constant unless we really have to; *normally use old constants when dropping universals.*[1] As before, the quantifier must begin the wff and we must replace the variable with the same constant throughout. So this next inference is fine:

(x)(Fx ⊃ Gx)
∴ (Fa ⊃ Ga)

This next example is wrong (because it drops the quantifier using two different constants):

(x)(Fx ⊃ Gx)
∴ (Fa ⊃ Gb)

This next example is also wrong (since the formula doesn't *begin* with a quantifier – instead it begins with a left-hand parenthesis – *drop only initial quantifiers*):

((x)Fx ⊃ (x)Gx)
∴ (Fa ⊃ Ga)

[1] Dropping a universal quantifier with a new letter assumes that something exists (or that our restricted universe of discourse is nonempty). Some systems (see §13.7) disallow this.

"((x)Fx ⊃ (x)Gx)" is an if-then and follows the if-then rules: if we have the first part "(x)Fx" true, we can get the second true; if we have the second part "(x)Gx" false, we can get the first false; if we get stuck, we make an assumption.

Here's an example of a proof:

All logicians are funny.
Someone is a logician.
∴ Someone is funny.

1 (x)(Lx ⊃ Fx) **Valid**
* 2 (∃x)Lx
[∴ (∃x)Fx
* 3 ┌ asm: ~(∃x)Fx
4 │ ∴ (x)~Fx {from 3}
5 │ ∴ La {from 2}
* 6 │ ∴ (La ⊃ Fa) {from 1}
7 │ **∴ Fa {from 5 and 6}**
8 └ **∴ ~Fa {from 4}**
9 ∴ (∃x)Fx {from 3; 7 contradicts 8}

For now, use the quantificational rules in this order:

- First reverse squiggles. We did this to get "(x)~Fx" in line 4.
- Then drop initial existentials, using a new constant each time. We did this to get "La" in line 5.
- Lastly, drop each initial universal once for each old constant. We did this to get "(La ⊃ Fa)" in line 6 and "~Fa" in line 8.

We starred lines 2, 3, and 6; starred lines largely can be ignored in deriving further lines. Star any wff on which you reverse squiggles or drop an existential:

* ~(x)Fx
———————
∴ (∃x)~Fx

* (∃x)Fx
———————
∴ Fa

Here the new line has the same information. *Don't star when dropping a* 0189 *universal*; we can never exhaust an "all" by deriving instances, and we may have to derive further things from it later.

Here's a simpler quantificational proof:

```
   1  (x)(Fx • Gx)    Valid
      [ ∴ (x)Fx
 * 2  ┌ asm: ~(x)Fx
 * 3  │ ∴ (∃x)~Fx    {from 2}
   4  │ ∴ ~Fa    {from 3}
   5  │ ∴ (Fa • Ga)    {from 1}
   6  └ ∴ Fa    {from 5}
   7 ∴ (x)Fx    {from 2; 4 contradicts 6}
```

Reverse squiggles to get "(∃x)~Fx" in line 3. Drop an existential to get "~Fa" in line 4. Then drop a universal to get "(Fa • Ga)" in line 5. Switching lines 4 and 5 would be wrong: if we drop the universal first using "a," then we can't drop the existential later using "a" (since then "a" would be old).

In doing proofs, first assume the conclusion's opposite; then use quantificational rules plus S- and I-rules to derive all you can. If you find a contradiction, apply RAA. If you're stuck and need to break a NOT-BOTH, OR, or IF-THEN, then make another assumption. If you get no contradiction and yet can't do anything further, then try to refute the argument. Here's a fuller statement of our strategy's quantificational steps:

1. FIRST REVERSE SQUIGGLES: For each unstarred, not-blocked-off line that begins with "~" and then a quantifier, derive a line using the reverse-squiggle rules. Star the original line.

2. THEN DROP EXISTENTIALS: For each unstarred, not-blocked-off line that begins with an existential quantifier, derive an instance using the next available new constant (but don't drop an existential if you already have a not-blocked-off instance in previous lines – so don't drop "(∃x)Fx" if you already have "Fc"). Star the original line.

3. LASTLY DROP UNIVERSALS: For each not-blocked-off line that begins with a universal quantifier, derive instances using each old constant. Don't star the original line; you may have to use it again. (Drop a universal using a new constant only if you've done everything else possible, making further assumptions if needed, and still have no old constants.)

Drop existentials before universals. Introduce a new constant each time you drop an existential, and use the same old constants when you drop a universal. And drop only initial quantifiers.

8.2a Exercise: LogiCola IEV

Prove each of these arguments to be valid (all are valid). 0190

~(∃x)Fx
∴ (x)~(Fx • Gx)

* 1 ~(∃x)Fx **Valid**
[∴ (x)~(Fx • Gx)
* 2 ┌ asm: ~(x)~(Fx • Gx)
* 3 │ ∴ (∃x)(Fx • Gx) {from 2}
4 │ ∴ (x)~Fx {from 1}
5 │ ∴ (Fa • Ga) {from 3}
6 │ **∴ ~Fa {from 4}**
7 └ **∴ Fa {from 5}**
8 ∴ (x)~(Fx • Gx) {from 2; 6 contradicts 7}

1. (x)Fx
∴ (x)(Gx ∨ Fx)

2. ~(∃x)(Fx • ~Gx)
∴ (x)(Fx ⊃ Gx)

3. ~(∃x)(Fx • Gx)
(∃x)Fx
∴ (∃x)~Gx

4. (x)((Fx ∨ Gx) ⊃ Hx)
∴ (x)(~Hx ⊃ ~Fx)

5. (x)(Fx ⊃ Gx)
(∃x)Fx
∴ (∃x)(Fx • Gx)

6. (x)(Fx ∨ Gx)
~(x)Fx
∴ (∃x)Gx

7. (x)~(Fx ∨ Gx)
∴ (x)~Fx

8. (x)(Fx ⊃ Gx)
(x)(Fx ⊃ ~Gx)
∴ (x)~Fx

9. $(x)(Fx \supset Gx)$
$(x)(\sim Fx \supset Hx)$
$\therefore (x)(Gx \vee Hx)$

10. $(x)(Fx \equiv Gx)$
$(\exists x)\sim Gx$
$\therefore (\exists x)\sim Fx$

8.2b Exercise: LogiCola IEV

First appraise intuitively. Then translate into logic (using the letters given) and prove to be valid (all are valid).

1. All who deliberate about alternatives believe in free will (at least implicitly).
All deliberate about alternatives.
∴ All believe in free will. [Use Dx and Bx; from William James.]

2. Everyone makes mistakes.
∴ Every logic teacher makes mistakes. [Use Mx and Lx.]

3. No feeling of pain is publicly observable.
All chemical processes are publicly observable.
∴ No feeling of pain is a chemical process. [Use Fx, Ox, and Cx. This attacks a form of materialism that identifies mental events with material events. We also could test this argument using syllogistic logic (Chapter 2).]

4. All (in the electoral college) who do their jobs are useless.
All (in the electoral college) who don't do their jobs are dangerous.
∴ All (in the electoral college) are useless or dangerous. [Use Jx for "x does their job," Ux for "x is useless," and Dx for "x is dangerous." Use the universe of discourse of electoral college members: take "(x)" to mean "for every electoral college member x" and don't translate "in the electoral college."] 0191

5. All that's known is experienced through the senses.
Nothing that's experienced through the senses is known.
∴ Nothing is known. [Use Kx and Ex. Empiricism (premise 1) plus skepticism about the senses (premise 2) yields general skepticism.]

6. No pure water is burnable.
Some Cuyahoga River water is burnable.
∴ Some Cuyahoga River water isn't pure water. [Use Px, Bx, and Cx. The Cuyahoga is a river in Cleveland that used to catch fire.]

7. Everyone who isn't with me is against me.
∴ Everyone who isn't against me is with me.   [Use Wx and Ax. These claims from the Gospels are sometimes thought to be incompatible.]

8. All basic laws depend on God's will.
∴ All basic laws about morality depend on God's will.   [Bx, Dx, Mx]

9. Some lies in unusual circumstances aren't wrong.
∴ Not all lies are wrong.   [Lx, Ux, Wx]

10. Nothing based on sense experience is certain.
Some logical inferences are certain.
All certain things are truths of reason.
∴ Some truths of reason are certain and aren't based on sense experience.   [Bx, Cx, Lx, Rx]

11. No truth by itself motivates us to action.
Every categorical imperative would by itself motivate us to action.
Every categorical imperative would be a truth.
∴ There are no categorical imperatives.   [Use Tx, Mx, and Cx. Immanuel Kant claimed that commonsense morality accepts categorical imperatives (objectively true moral judgments that command us to act and that we must follow if we are to be rational); but some thinkers argue against the idea.]

12. Every genuine truth claim is either experimentally testable or true by definition.
No moral judgments are experimentally testable.
No moral judgments are true by definition.
∴ No moral judgments are genuine truth claims.   [Use Gx, Ex, Dx, and Mx. This is logical positivism's argument against moral truths.]

13. Everyone who can think clearly would do well in logic.
Everyone who would do well in logic ought to study logic.
Everyone who can't think clearly ought to study logic.
∴ Everyone ought to study logic.   [Tx, Wx, Ox]

8.3 Easier refutations

Applying our proof strategy to an invalid argument leads to a refutation: 0192

Someone is short.
Someone is tall.
∴ Someone is both short and tall.

```
*  1  (∃x)Sx    Invalid
*  2  (∃x)Tx
      [ ∴ (∃x)(Sx • Tx)
*  3  asm: ~(∃x)(Sx • Tx)
   4  ∴ (x)~(Sx • Tx)    {from 3}
   5  ∴ Sa    {from 1}
   6  ∴ Tb    {from 2}
*  7  ∴ ~(Sa • Ta)    {from 4}
*  8  ∴ ~(Sb • Tb)    {from 4}
   9  ∴ ~Ta    {from 5 and 7}
  10  ∴ ~Sb    {from 6 and 8}
```

a, b

Sa, ~Ta Tb, ~Sb

Reverse a squiggle (line 4). Drop two existentials, using a new constant each time (lines 5 and 6). Drop the universal twice, using "a" and "b" (lines 7 and 8). Getting no contradiction, we gather simple wffs for a *refutation* (here a "simple wff" is one containing only capital letters, zero or more constants, and zero or one squiggles). We get a little possible world with two people, a and b, where a is short and not tall, but b is tall and not short. The argument is invalid, since this possible world makes the premises all true (someone is short and someone is tall) but the conclusion false (no one is both short and tall).

If we try to prove an invalid argument, we'll instead be led to a refutation – a little possible world with various individuals (like a and b) and simple truths about them (like Sa and ~Sb) that make the premises all true and conclusion false. In evaluating premises and conclusion, use these rules to evaluate each formula or subformula that starts with a quantifier:

An *existential* wff is true if and only if *at least one case* is true.
A *universal* wff is true if and only if *all cases* are true.

Premise "(∃x)Sx" is true because at least one case ("Sa") is true, and premise "(∃x)Tx" is true because at least one case ("Tb") is true.[1] But conclusion

[1] SOME is like OR: something holds in this case OR that case OR that case … – so a single true

"(∃x)(Sx • Tx)" is false because both cases are false:

(Sa • Ta) = (1 • 0) = 0

(Sb • Tb) = (0 • 1) = 0

Always check that your refutation works. If you don't get premises all 1 and conclusion 0, then you did something wrong; look at what you did with the wff that came out wrong (a premise that's 0 or ?, or a conclusion that's 1 or ?).

These two rules are crucial for working out proofs and refutations: 0193

- For each initial existential quantifier, introduce a new constant.
- For each initial universal quantifier, derive an instance for each old constant.

If you have two existentials, don't drop both using the same constant – and don't drop just one existential. And if you have two constants, then drop any universals using *both* constants; if in our example we dropped the universal in "(x)∼(Sx • Tx)" using "a" but not "b," then our refutation would fail:

a, b

Sa, ∼Ta, Tb

a is short and not tall, b is tall

Since "Sb" is unknown, our conclusion "(∃x)(Sx • Tx)" would also be unknown (because the second case with "b" is unknown):

(Sa • Ta) = (1 • 0) = 0

(**Sb** • Tb) = (**?** • 1) = **?**

The "Someone is both short and tall" conclusion is unknown, since our world doesn't exclude b being short (besides being tall). We avoid such problems if we drop each initial universal quantifier using each old constant; here we'd go from "(x)∼(Sx • Tx)" to "∼(Sb • Tb)," which would lead to "∼Sb."

As we refute arguments, we'll often have to evaluate premises or conclusions that don't *start* with quantifiers, such as these wffs:

case makes a SOME true. ALL is like AND: something holds in this case AND that case AND that case … – so a single false case makes an ALL false.

~(x)Sx
~(x)Sx

~(x)(Sx ∨ Tx)
~(x)(Sx ∨ Tx)

~(∃x)(Sx • Tx)
~(∃x)(Sx • Tx)

Identify any *subformulas* that start with quantifiers (as highlighted here). Evaluate each subformula to be 1 or 0, and then apply "~" to reverse the result. On our short-tall refutation, "(x)Sx" = 0 and so "~(x)Sx" = 1. Likewise, "(x)(Sx ∨ Tx)" = 1, and so "~(x)(Sx ∨ Tx)" = 0; and "(∃x)(Sx • Tx)" = 0, and so "~(∃x)(Sx • Tx)" = 1. In evaluating a wff that starts with a squiggle and then a quantifier, evaluate the wff without the squiggle and then give the original wff the opposite value. Divide and conquer!

Possible worlds for refutations must contain at least one entity. We seldom need more than two entities.

8.3a Exercise: LogiCola IEI

Prove each of these arguments to be invalid (all are invalid). 0194

~(x)(Fx ∨ Gx)
∴ ~(∃x)Gx

* 1 ~(x)(Fx ∨ Gx) **Invalid**
[∴ ~(∃x)Gx
* 2 asm: (∃x)Gx
* 3 ∴ (∃x)~(Fx ∨ Gx) {from 1}
* 4 ∴ ~(Fa ∨ Ga) {from 3}
5 ∴ ~Fa {from 4}
6 ∴ ~Ga {from 4}
7 ∴ Gb {from 2}

a, b

~Fa, ~Ga, Gb

1. (∃x)Fx
∴ (x)Fx

2. (∃x)Fx
(∃x)Gx
∴ (∃x)(Fx • Gx)

3. (∃x)(Fx ∨ Gx)
~(x)Fx
∴ (∃x)Gx

4. (∃x)Fx
∴ (∃x)~Fx

5. ~(∃x)(Fx • Gx)
(x)~Fx
∴ (x)Gx

6. (x)(Fx ⊃ Gx)
~(x)Gx
∴ (x)~(Fx • Gx)

7. (x)((Fx • Gx) ⊃ Hx)
(∃x)Fx
(∃x)Gx
∴ (∃x)Hx

8. (∃x)(Fx ∨ ~Gx)
(x)(~Gx ⊃ Hx)
(∃x)(Fx ⊃ Hx)
∴ (∃x)Hx

9. (∃x)~(Fx ∨ Gx)
(∃x)Hx
~(∃x)Fx
∴ ~(x)(Hx ⊃ Gx)

10. (∃x)~Fx
(∃x)~Gx
∴ (∃x)(Fx ≡ Gx)

8.3b Exercise: LogiCola IEC

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation).

1. Some butlers are guilty.
∴ All butlers are guilty. [Use Bx and Gx.]

2. No material thing is infinite.
Not everything is material.
∴ Something is infinite. [Use Mx and Ix.]

3. Some smoke.
Not all have clean lungs.
∴ Some who smoke don't have clean lungs. [Use Sx and Cx.]

4. Some Marxists plot violent revolution.
Some faculty members are Marxists.
∴ Some faculty members plot violent revolution. [Mx, Px, Fx] 0195

5. All valid arguments that have "ought" in the conclusion also have "ought" in the premises.
All arguments that seek to deduce an "ought" from an "is" have "ought" in the conclusion but don't have "ought" in the premises.
∴ No argument that seeks to deduce an "ought" from an "is" is valid. [Use Vx for "x is valid," Cx for "x has 'ought' in the conclusion," Px for "x has 'ought' in the premises," Dx for "x seeks to deduce an 'ought' from an 'is,'" and the universe of discourse of arguments. This one is difficult to translate.]

6. Every kick returner who is successful is fast.
∴ Every kick returner who is fast is successful. [Kx, Sx, Fx]

7. All exceptionless duties are based on the categorical imperative.
All non-exceptionless duties are based on the categorical imperative.
∴ All duties are based on the categorical imperative. [Use Ex, Bx, and the universe of discourse of duties; from Kant, who based all duties on his supreme moral principle, called "the categorical imperative."]

8. All who aren't crazy agree with me.
∴ No one who is crazy agrees with me. [Cx, Ax]

9. Everything can be conceived.
Everything that can be conceived is mental.
∴ Everything is mental. [Use Cx and Mx; from George Berkeley, who attacked materialism by arguing that everything is mental and that matter doesn't exist apart from mental sensations; so a chair is just a collection of experiences. Bertrand Russell thought premise 2 was confused.]

10. All sound arguments are valid.
∴ All invalid arguments are unsound. [Use Sx and Vx and the universe of discourse of arguments.]

11. All trespassers are eaten.
∴ Some trespassers are eaten. [Use Tx and Ex. The premise is from a sign on the Appalachian Trail in northern Virginia. Traditional logic (§2.8) takes "all A is B" to entail "some A is B"; modern logic takes "all A is B" to mean "whatever is A also is B" – which can be true even if there are no A's.]

12. Some necessary being exists.
All necessary beings are perfect beings.
∴ Some perfect being exists. [Use Nx and Px. Kant claimed that the cosmological argument for God's existence at most proves premise 1; it doesn't prove the existence of a perfect God unless we add premise 2. But premise 2, by the next argument, presupposes the central claim of the ontological argument – that some perfect being is a necessary being. So, Kant claimed, the cosmological argument presupposes the ontological argument.]

13. All necessary beings are perfect beings.
∴ Some perfect being is a necessary being. [Use Nx and Px. Kant followed traditional logic (see problem 11) in taking "all A is B" to entail "some A is B."] 0196

14. No one who isn't a logical positivist holds the verifiability criterion of meaning.
∴ All who hold the verifiability criterion of meaning are logical positivists. [Use Lx and Hx. The verifiability criterion of meaning says that every genuine truth claim is either experimentally testable or true by definition.]

15. No pure water is burnable.
Some Cuyahoga River water isn't burnable.
∴ Some Cuyahoga River water is pure water. [Use Px, Bx, and Cx.]

8.4 Harder translations

We'll now start using statement letters (like "S" for "It's snowing") and individual constants (like "r" for "Romeo"); here's an example:

If it's snowing, then Romeo is cold = (S ⊃ Cr)

Here "S," since it's a capital letter not followed by a small letter, represents a whole statement. And "r," since it's a small letter between "a" and "w," is a constant that stands for a specific person or thing.

We'll also start using multiple and non-initial quantifiers. From now on, use this expanded rule about what quantifier to use and where to put it:

Where the English has "**all**" or "**every**," put this in the wff: "**(x)**."

Where the English has "**not all**" or "**not every**," put this in the wff: "**$\sim$(x)**."

Where the English has "**some**," put this in the wff: "**($\exists$x)**."

Where the English has "**no**," put this in the wff: "**$\sim$($\exists$x)**."

If all are Italian, then Romeo is Italian = $((x)Ix \supset Ir)$

Since "if" translates as "(," likewise "if all" translates as "((x)." As you translate, mimic the English word order:

all not = $(x)\sim$
not all = $\sim(x)$

all either = $(x)($
either all = $((x)$

if all either = $((x)($
if either all = $(((x)$

Use a separate quantifier for each "all," "some," and "no":

If all are Italian, then all are lovers
= $((x)Ix \supset (x)Lx)$

If not everyone is Italian, then some aren't lovers
= $(\sim(x)Ix \supset (\exists x)\sim Lx)$

If no Italians are lovers, then some Italians are not lovers
= $(\sim(\exists x)(Ix \bullet Lx) \supset (\exists x)(Ix \bullet \sim Lx))$ 0197

"Any" differs in subtle ways from "all" (which translates into a "(x)" that mirrors where "all" occurs in the English sentence). "Any" has two different but equivalent translation rules; here's the easier rule, with examples:

To translate "any," first rephrase the sentence so it means the same thing but doesn't use "any"; then translate the second sentence.

"Not any …" = "No …."
"If any …" = "If some …."
"Any …" = "All …."

Not anyone is rich = No one is rich
= $\sim(\exists x)Rx$

Not any Italian is a lover = No Italian is a lover
= $\sim(\exists x)(Ix \bullet Lx)$

If anyone is just, there will be peace = If someone is just, there will be peace
= $((\exists x)Jx \supset P)$

Our second rule usually gives a formula that's different but equivalent:

To translate "any," put a "(x)" at the *beginning* of the wff, regardless of where the "any" occurs in the sentence.

Not anyone is rich = For all x, x isn't rich
= $(x)\sim Rx$

Not any Italian is a lover = For all x, x isn't both Italian and a lover
= $(x)\sim(Ix \bullet Lx)$ ⇐ Note "•" here!

If anyone is just, there will be peace = For all x, if x is just there will be peace
= $(x)(Jx \supset P)$

"Any" at the beginning of a sentence usually just means "all." So "Any Italian is a lover" means "All Italians are lovers."

8.4a Exercise: LogiCola H (HM & HT)

Translate these English sentences into wffs. Recall that our translation rules are rough guides and sometimes don't work; so read your formula carefully to make sure it reflects what the English means.

If everyone is evil, then Gensler is evil.
$((x)Ex \supset Eg)$

1. Gensler is either crazy or evil.
2. If Gensler is a logician, then some logicians are evil.
3. If everyone is a logician, then everyone is evil.
4. If all logicians are evil, then some logicians are evil.
5. If someone is evil, it will rain.

6. If everyone is evil, it will rain.
7. If anyone is evil, it will rain. 0198
8. If Gensler is a logician, then someone is a logician.
9. If no one is evil, then no one is an evil logician.
10. If all are evil, then all logicians are evil.
11. If some are logicians, then some are evil.
12. All crazy logicians are evil.
13. Everyone who isn't a logician is evil.
14. Not everyone is evil.
15. Not anyone is evil.
16. If Gensler is a logician, then he's evil.
17. If anyone is a logician, then Gensler is a logician.
18. If someone is a logician, then he or she is evil.
19. Everyone is an evil logician.
20. Not any logician is evil.

8.5 Harder proofs

Now we come to proofs using formulas with multiple or non-initial quantifiers. Such proofs, while needing no new inference rules, are often tricky and require multiple assumptions. As before, drop only initial quantifiers:

Both of these are wrong:

((x)Fx ⊃ (x)Gx)
∴ (Fa ⊃ (x)Gx)

((x)Fx ⊃ (x)Gx)
∴ (Fa ⊃ Ga)

"((x)Fx ⊃ (x)Gx)" is an if-then; to infer with it, we need the first part true or the second part false – as in these examples:

Both of these are right:

((x)Fx ⊃ (x)Gx)
(x)Fx

∴ (x)Gx

((x)Fx ⊃ (x)Gx)
~(x)Gx
∴ ~(x)Fx

If we get stuck, we may need to assume one side or its negation.

Here's a proof using a formula with multiple quantifiers: 0199

If *some* are enslaved, then all have their freedom threatened.
∴ If this person is enslaved, then I have my freedom threatened.

* 1 ((∃x)Sx ⊃ (x)Tx) **Valid**
[∴ (St ⊃ Ti)
* 2 asm: ~(St ⊃ Ti)
3 **∴ St {from 2}**
4 **∴ ~Ti {from 2}**
5 asm: ~(∃x)Sx {break 1}
6 ∴ (x)~Sx {from 5}
7 **∴ ~St {from 6}**
8 ∴ (∃x)Sx {from 5; 3 contradicts 7}
9 ∴ (x)Tx {from 1 and 8}
10 **∴ Ti {from 9}**
11 ∴ (St ⊃ Ti) {from 2; 4 contradicts 10}

After the assumption, we apply an S-rule to get lines 3 and 4. Then we're stuck, since we can't drop the non-initial quantifiers in 1. So we make a second assumption in line 5, get a contradiction, and derive 8. We soon get a second contradiction to complete the proof.

Here's an invalid argument:

If *all* are enslaved, then all have their freedom threatened.
∴ If this person is enslaved, then I have my freedom threatened.

1 ((x)Sx ⊃ (x)Tx) **Invalid**
[∴ (St ⊃ Ti)
* 2 asm: ~(St ⊃ Ti)
3 ∴ St {from 2}
4 ∴ ~Ti {from 2}
** 5 asm: ~(x)Sx {break 1}
** 6 ∴ (∃x)~Sx {from 5}
7 ∴ ~Sa {from 6}

t, i, a

St, ∼Ti, ∼Sa

In evaluating the premise, first identity and evaluate *subformulas* that start with quantifiers (these are highlighted here), and then plug in 1 or 0 for these:

For "((x)Sx ⊃ (x)Tx)," we first evaluate "(x)Sx" and "(x)Tx":
"(x)Sx" is false because "Sa" is false.
"(x)Tx" is false because "Ti" is false.
Replace both with "0."
We get "(0 ⊃ 0)," which simplifies to "1."
So "((x)Sx ⊃ (x)Tx)" is true.

So the premise is true. Since the conclusion is false, the argument is invalid.

As we refute invalid arguments, we'll often have complex premises or conclusions to evaluate, such as these wffs:

((x)Sx ⊃ (x)Tx)
((x)Sx ⊃ (x)Tx)

((x)(Fx ∨ Gx) ⊃ (∼(∃x)Gx • (∃x)∼Hx))
((x)(Fx ∨ Gx) ⊃ (∼(∃x)Gx • (∃x)∼Hx))

(∼(x)(Fx ⊃ Gx) ≡ ∼(∃x)Hx)
(∼(x)(Fx ⊃ Gx) ≡ ∼(∃x)Hx)

Identity any *subformulas* that start with quantifiers (as highlighted 0200 here). Evaluate each such subformula to be 1 or 0, replace it with 1 or 0, and figure out whether the whole formula is 1 or 0. Divide and conquer!

8.5a Exercise: LogiCola I (HC & MC)

Say whether each is valid (and give a proof) or invalid (and give a refutation).

(x)(Mx ∨ Fx)
∴ ((x)Mx ∨ (x)Fx)

(This is like arguing that, since everyone is male or female, thus either everyone is male or everyone is female.)

1 (x)(Mx ∨ Fx) **Invalid**
[∴ ((x)Mx ∨ (x)Fx)
* 2 asm: ~((x)Mx ∨ (x)Fx)
* 3 ∴ ~(x)Mx {from 2}
* 4 ∴ ~(x)Fx {from 2}
* 5 ∴ (∃x)~Mx {from 3}
* 6 ∴ (∃x)~Fx {from 4}
7 ∴ ~Ma {from 5}
8 ∴ ~Fb {from 6}
* 9 ∴ (Ma ∨ Fa) {from 1}
* 10 ∴ (Mb ∨ Fb) {from 1}
11 ∴ Fa {from 7 and 9}
12 ∴ Mb {from 8 and 10}

a, b

Fa, ~Ma
Mb, ~Fb

1. (x)(Fx ∨ Gx)
~Fa
∴ (∃x)Gx

2. (x)(Ex ⊃ R)
∴ ((∃x)Ex ⊃ R)

3. ((x)Ex ⊃ R)
∴ (x)(Ex ⊃ R)

4. ((∃x)Fx ∨ (∃x)Gx)
∴ (∃x)(Fx ∨ Gx)

5. ((∃x)Fx ⊃ (∃x)Gx)
∴ (x)(Fx ⊃ Gx)

6. (x)((Fx ∨ Gx) ⊃ Hx)
Fm
∴ Hm

7. Fj
(∃x)Gx
(x)((Fx • Gx) ⊃ Hx)
∴ (∃x)Hx

8. ((∃x)Fx ⊃ (x)Gx)
~Gp
∴ ~Fp

9. (∃x)(Fx ∨ Gx)
∴ ((x)~Gx ⊃ (∃x)Fx)

10. ~(∃x)(Fx • Gx)
~Fd
∴ Gd

11. (x)(Ex ⊃ R)
∴ ((x)Ex ⊃ R)

12. (x)(Fx • Gx)
∴ ((x)Fx • (x)Gx)

13. (R ⊃ (x)Ex)
∴ (x)(R ⊃ Ex)

14. ((x)Fx ∨ (x)Gx)
∴ (x)(Fx ∨ Gx)

15. ((∃x)Ex ⊃ R)
∴ (x)(Ex ⊃ R)

8.5b Exercise: LogiCola I (HC & MC)

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation).

1. Everything has a cause.
If the world has a cause, then there is a God.
∴ There is a God. [Use Cx for "x has a cause," w for "the world," and G for "There is a God" (which we needn't here break down into "(∃x)Gx" – "For some x, x is a God"). A student of mine suggested this argument; but the next example shows that premise 1 can as easily lead to the opposite conclusion.] 0201

2. Everything has a cause.
If there is a God, then something doesn't have a cause (namely, God).
∴ There is no God. [Use Cx and G. The next example qualifies "Everything has a cause" to avoid the problem; some prefer an argument based on "Every *contingent being or set of such beings* has a cause."]

3. Everything that began to exist has a cause.
The world began to exist.
If the world has a cause, then there is a God.
∴ There is a God. [Use Bx, Cx, w, and G. This "Kalam argument" is from William Craig and James Moreland; they defend premise 2 by appealing to the Big Bang theory, the law of entropy, and the impossibility of an actual infinite.]

4. If everyone litters, then the world will be dirty.
∴ If you litter, then the world will be dirty. [Lx, D, u]

5. Anything enjoyable is either immoral or fattening.
∴ If nothing is immoral, then everything that isn't fattening isn't enjoyable. [Ex, Ix, Fx]

6. Anything that can be explained either can be explained as caused by scientific laws or can be explained as resulting from a free choice of a rational being.
The totality of basic scientific laws can't be explained as caused by scientific laws (since this would be circular).
∴ Either the totality of basic scientific laws can't be explained or else it can be explained as resulting from a free choice of a rational being (God). [Use Ex for "x can be explained," Sx for "x can be explained as caused by scientific laws," Fx for "x can be explained as resulting from a free choice of a rational being," and t for "the totality of scientific laws." This one is from R. G. Swinburne.]

7. If someone knows the future, then no one has free will.
∴ No one who knows the future has free will. [Kx, Fx]

8. If everyone teaches philosophy, then everyone will starve.
∴ Everyone who teaches philosophy will starve. [Tx, Sx]

9. No proposition based on sense experience is logically necessary.
∴ Either no mathematical proposition is based on sense experience, or no mathematical proposition is logically necessary. [Use Sx, Nx, and Mx, and the universe of propositions; from the logical positivist A. J. Ayer.]

10. Any basic social rule that people would agree to if they were free and rational but ignorant of their place in society (whether rich or poor, white or black, male or female) is a principle of justice.
The equal-liberty principle and difference principle are basic social rules that people would agree to if they were free and rational but ignorant of their place in society.
∴ The equal-liberty principle and difference principle are principles of justice. [Use Ax, Px, e, and d; from John Rawls. *Equal-liberty* says that everyone is entitled to the greatest liberty compatible with an equal liberty for all others; *difference* says that wealth is to be distributed equally, except for inequalities that provide incentives that ultimately benefit everyone and are equally open to all.] 0202

11. If there are no necessary beings, then there are no contingent beings.
∴ All contingent beings are necessary beings. [Use Nx and Cx. Aquinas accepted the premise but not the conclusion.]

12. Anything not disproved that's of practical value to one's life to believe ought to be believed.
Free will isn't disproved.
∴ If free will is of practical value to one's life to believe, then it ought to be believed. [Use Dx, Vx, Ox, f (for "free will"), and the universe of discourse of beliefs; from William James.]

13. If the world had no temporal beginning, then some series of moments before the present moment is a completed infinite series.
There's no completed infinite series.
∴ The world had a temporal beginning. [Use Tx for "x had a temporal beginning," w for "the world," Mx for "x is a series of moments before the present moment," and Ix for "x is a completed infinite series." This one and the next are from Immanuel Kant, who thought our intuitive metaphysical principles lead to conflicting conclusions and thus can't be trusted.]

14. Everything that had a temporal beginning was caused to exist by something previously in existence.
If the world was caused to exist by something previously in existence, then there was time before the world began.
If the world had a temporal beginning, then there was no time before the world began.
∴ The world didn't have a temporal beginning. [Use Tx for "x had a temporal beginning," Cx for "x was caused to exist by something previously in existence," w for "the world," and B for "There was time before the world began."]

15. If emotivism is true, then “X is good” means “Hurrah for X!” and all moral judgments are exclamations.
All exclamations are inherently emotional.
“This dishonest income tax exemption is wrong” is a moral judgment.
“This dishonest income tax exemption is wrong” isn’t inherently emotional.
∴ Emotivism isn’t true. [T, H, Mx, Ex, Ix, t]

16. If everything is material, then all prime numbers are composed of physical particles.
Seven is a prime number.
Seven isn’t composed of physical particles.
∴ Not everything is material. [Mx, Px, Cx, s]

17. If everyone lies, the results will be disastrous.
∴ If anyone lies, the results will be disastrous. [Lx, D] 0203

18. Everyone makes moral judgments.
Moral judgments logically presuppose beliefs about God.
If moral judgments logically presuppose beliefs about God, then everyone who makes moral judgments believes (at least implicitly) that there is a God.
∴ Everyone believes (at least implicitly) that there is a God. [Use Mx for “x makes moral judgments,” L for “Moral judgments logically presuppose beliefs about God,” and Bx for “x believes (at least implicitly) that there is a God.” This is from the Jesuit theologian Karl Rahner.]

19. “x = x” is a basic law.
“x = x” is true in itself, and not true because someone made it true.
If “x = x” depends on God’s will, then “x = x” is true because someone made it true.
∴ Some basic laws don’t depend on God’s will. [Use e (for “x = x”), Bx, Tx, Mx, and Dx.]

20. Nothing that isn’t caused can be integrated into the unity of our experience.
Everything that we could experientially know can be integrated into the unity of our experience.
∴ Everything that we could experientially know is caused. [Use Cx, Ix, and Ex; from Immanuel Kant. The conclusion is limited to objects of possible experience – since it says “Everything *that we could experientially know* is caused”; Kant thought the unqualified “Everything is caused” leads to contradictions (see # 1 and 2).]

21. If everyone deliberates about alternatives, then everyone believes (implicitly) in free will.
∴ All who deliberate about alternatives believe (implicitly) in free will. [Dx, Bx]

22. All who are consistent and think that abortion is normally permissible will consent to the idea of their having been aborted in normal circumstances.
You don't consent to the idea of your having been aborted in normal circumstances.
∴ If you're consistent, then you won't think that abortion is normally permissible.
[Use Cx, Px, Ix, and u. See my article in January 1986 *Philosophical Studies* or the synthesis chapter of my *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018).]

8.6 Copi proofs

We earlier discussed the traditional Copi proof method for propositional logic (§7.5). This method can also be used for quantificational logic.

For each quantifier (universal and existential), Copi has *instantiation rules* (to drop a quantifier) and *generalization rules* (to add a quantifier). Existential instantiation (EI) is the same as our drop-existential rule (DE). It holds regardless of what variable replaces "x," what constant replaces "a," and what wffs replace "Fx" / "Fa" – provided that the two wffs are identical except that wherever the 0204 variable occurs freely[1] in the former the same constant occurs in the latter:

EI Existential instantiation
(∃x)Fx → Fa,
use a *new constant*

Here the constant must not have occurred in any previous step of the proof or in the original conclusion. As before, "→" in all these rules means that we can infer whole lines from left to right. Existential generalization works the opposite way, and isn't subject to the restriction that the constant has to be new:

EG Existential generalization
Fa → (∃x)Fx

Another form of the rule uses a variable in place of "a":

[1] An instance of a variable occurs *freely* in a formula if it's not part of a wff that begins with a quantifier using that variable; just the first instance of "x" in "(Fx • (x)Gx)" occurs freely.

EG Existential generalization
Fy → (∃x)Fx

This form holds regardless of what variables replace "x" and "y" and what wffs replace "Fx" / "Fy" – provided that the two wffs are identical except that wherever the variable that replaces "x" occurs freely in the former the variable that replaces "y" occurs freely in the latter.

Universal instantiation (UI) is like our drop-universal rule (DE), except that it also has two forms. The first form holds regardless of what variable replaces "x," what constant replaces "a," and what wffs replace "Fx" / "Fa" – provided that the two wffs are identical except that wherever the variable occurs freely in the former the same constant occurs in the latter:

UI Universal instantiation
(x)Fx → Fa

Here the constant needn't be new, and so it could have occurred earlier in the proof or in the original conclusion. And again a second form uses a variable in place of "a." This form holds regardless of what variables replace "x" and "y" and what wffs replace "Fx" / "Fy" – provided that the two wffs are identical except that wherever the variable that replaces "x" occurs freely in the former the variable that replaces "y" occurs freely in the latter:

UI Universal instantiation
(x)Fx → Fy

0205 Universal generalization works the opposite way:

UG Universal generalization
Fy → (x)Fx, "y" can't occur in an assumption

Here the variable that replaces "y" can't occur in an assumption.

Copi has a replacement rule much like our reverse-squiggle rules (any variable can uniformly replace "x" and any wff can uniformly replace "P"):

QN Quantifier negation
(x)P = ~(∃x)~P (∃x)P = ~(x)~P

This lets us switch, for example, one instance of "(x)Fx" and "~(∃x)~Fx" anywhere in a wff. I'll take the Copi proof system for quantificational logic to include these five rules: EI, EG, UI, UG, and QN.

To see how these work, I'll now prove the valid arguments in this chapter's explanation sections. Here's a Copi proof for the first example in §8.2:

Conclusion: (∃x)Fx

1 (x)(Lx ⊃ Fx)
2 (∃x)Lx
3 La {EI 2}
4 (La ⊃ Fa) {UI 1}
5 Fa {MP 3+4}
6 (∃x)Fx {EG 5}

And here's a Copi proof for the second example in §8.2:

Conclusion: (x)Fx

1 (x)(Fx • Gx)
2 (Fy • Gy) {UI 1}
3 Fy {SM 2}
4 (x)Fx {UG 3}

And here's a Copi proof for the valid example in §8.5:

Conclusion: (St ⊃ Ti)

1 ((∃x)Sx ⊃ (x)Tx)
2 St {Assume} *
3 (∃x)Sx {EG 2} *
4 (x)Tx {MP 1+3} *
5 Ti {UI 4} *
6 (St ⊃ Ti) {CP 2+5} 0206

If you've mastered Copi propositional proofs, the quantificational proofs won't be too difficult. If you're really confused on how to start, try assuming the opposite of the conclusion and deriving a contradiction. Again, use Copi proofs only on valid arguments; if you try the Copi procedure on an invalid argument, you won't derive the conclusion and you won't reach a natural "stopping point" that gives you a refutation of the argument's validity.

8.5a and 8.5b Exercise: LogiCola IEO

Do Copi proofs for problems in §§8.2a and 8.2b (all are valid). These are easier problems.

8.5c and 8.5d Exercise: LogiCola IHO and IMO

Do Copi proofs for problems in §8.5a (just the valid ones, namely 1, 2, 4, 6, 8, 9, 11, 12, 13, 14, and 15) and §8.5b (just the valid ones, namely 1, 2, 3, 5, 6, 7, 10, 12, 13, 14, 15, 16, 18, 19, 20, and 22). These are harder problems.

9 Relations and Identity

We now bring quantificational logic up to full power by adding **identity** and **relational statements**, like "a=b" and "Lrj" ("Romeo loves Juliet").

9.1 Identity translations

Our third rule for forming quantificational wffs introduces "=" ("equals"):

> The result of writing a small letter and then "=" and then a small letter is a wff.

This lets us construct wffs like these:

x=y = x equals y.
r=l = Romeo is the lover of Juliet.
~p=l = Paris isn't the lover of Juliet.

We negate an identity wff by writing "~" in front. Neither "r=l" nor "~p=l" use parentheses, since these aren't needed to avoid ambiguity.

The simplest use of "=" translates an "is" that goes between singular terms. Recall the difference between general and singular terms:

Use capital letters for **general terms**, which *describe* or put in a *category*: I = an Italian C = charming F = drives a Ford Use capitals for "a so and so," adjectives, and verbs.
Use small letters for **singular terms**, which pick out a *specific* person or thing: i = the richest Italian t = this child r = Romeo Use small letters for "the so and so," "this so and so," and proper names.

Compare these two forms: 0208

Predication Lr Romeo is a lover
Identity r=l Romeo is the lover of Juliet

Use "=" for "is" if both sides are singular terms (represented by small letters). The "is" of identity can be replaced with "is identical to" or "is the same entity as," and can be reversed (so if x=y then y=x).

We can translate "other than," "besides," and "alone" using identity:

Someone *other than* (*besides*) Romeo is rich
= $(\exists x)(\sim x=r \bullet Rx)$
For some x, x ≠ Romeo and x is rich

Romeo *alone* is rich
= $(Rr \bullet \sim(\exists x)(\sim x=r \bullet Rx))$
Romeo is rich and it's false that, for some x, x ≠ Romeo and x is rich

These translations also work if we switch conjuncts ("∼x=r" and "Rx") or switch the order of letters in an identity ("∼r=x" works in place of "∼x=r"). We also can translate some numerical notions, for example:

At least two are rich
= $(\exists x)(\exists y)(\sim x=y \bullet (Rx \bullet Ry))$
For some x and some y: x≠y, x is rich, and y is rich

The pair of quantifiers "(∃x)(∃y)" ("for some x and some y") doesn't say whether x and y are identical; so we need "∼x=y" to say they *aren't*. Henceforth we'll often need more variable letters than just "x" to keep references straight. In general, it doesn't matter which variable letters we use; we can translate "At least one is rich" as "(∃x)Rx" or "(∃y)Ry" or "(∃z)Rz."

We can express "exactly one" and "exactly two" (and "exactly n," for any specific whole number n):

Exactly one is dark

= (∃x)(Dx • ~(∃y)(~y=x • Dy))
For some x, x is dark and there's no y such that y≠x and y is dark

Exactly two are dark
= (∃x)(∃y)(((Dx • Dy) • ~x=y) • ~(∃z)((~z=x • ~z=y) • Dz))
For some x and some y, x is dark and y is dark and x≠y and there's no z such that z≠x and z≠y and z is dark

We also can express addition. Here's a Loglish paraphrase of "1 + 1 = 2" and the corresponding formula: 0209

If exactly one being is F and exactly one being is G and nothing is F-and-G, then exactly two beings are F-or-G.

((((∃x)(Fx • ~(∃y)(~y=x • Fy)) • (∃x)(Gx • ~(∃y)(~y=x • Gy))) • ~(∃x)(Fx • Gx)) ⊃ (∃x)(∃y)(((Fx ∨ Gx) • (Fy ∨ Gy)) • (~x=y • ~(∃z)((~z=x • ~z=y) • (Fz ∨ Gz)))))

We could prove our "1 + 1 = 2" formula by assuming its denial and deriving a contradiction. While this would be tedious, it's interesting that it could be done. In principle, we could prove "2 + 2 = 4" and "5 + 7 = 12" – and the additions on your income tax form. Some mean teachers assign such homework problems.

9.1a Exercise: LogiCola H (IM & IT)

Translate these English sentences into wffs.

Jim is the goalie and is a student.
(j=g • Sj)

1. Aristotle is a logician.
2. Aristotle is the greatest logician.
3. Aristotle isn't Plato.
4. Someone besides Aristotle is a logician.
5. There are at least two logicians.
6. Aristotle alone is a logician.
7. All logicians other than Aristotle are evil.

8. No one besides Aristotle is evil.
9. The philosopher is Aristotle.
10. There's exactly one logician.
11. There's exactly one evil logician.
12. Everyone besides Aristotle and Plato is evil.
13. If the thief is intelligent, then you aren't the thief.
14. Carol is my only sister.
15. Alice runs but isn't the fastest runner.
16. There's at most one king.
17. The king is bald.
18. There's exactly one king and he is bald.

9.2 Identity proofs

We need two new inference rules for identity. This self-identity rule holds regardless of what constant replaces "a":

Self-identity SI
a=a

0210 This is an **axiom** – a basic unproved assertion that can be used to prove other things. This rule says that we may assert a self-identity as a "derived line" anywhere in a proof, regardless of earlier lines. Adding "a=a" can be useful if we already have "∼a=a" (since then we get a contradiction) or already have a line like "(a=a ⊃ Gb)" (since then we can apply an I-rule). Our self-identity line will mention this previous line; it might say "∴ b=b {self-identity, to contradict 3}."

This substitute-equals rule is based on interchangeability of identicals: if a=b, then whatever is true of a is true of b, and vice versa. This rule holds regardless of what constants replace "a" and "b" and what wffs replace "Fa" and "Fb" – provided that the two wffs are alike except that the constants are interchanged in one or more occurrences:

Substitute Equals SE
a=b, Fa → Fb

Here's an easy identity proof:

I weigh 180 pounds.
My mind doesn't weigh 180 pounds.
∴ I'm not identical to my mind.

1 Wi **Valid**
2 **~Wm**
[∴ ~i=m
3 ┌ asm: i=m
4 └ **∴ Wm {from 1 and 3}**
5 ∴ ~i=m {from 3; 2 contradicts 4}

Line 4 follows by substituting equals; if i and m are identical, then whatever is true of one is true of the other.

Here's an easy invalid argument and its refutation:

The bankrobber wears size-twelve shoes.
You wear size-twelve shoes.
∴ You're the bankrobber.

1 Wb **Invalid**
2 Wu
[∴ u=b
3 asm: ~u=b

b, u

Wb, Wu, ~u=b

Since we can't infer anything here (we can't do much with "~u=b"), we set up a possible world to refute the argument. This world contains two distinct persons, the bankrobber and you, each wearing size-twelve shoes. Since the premises are all true and conclusion false in this world, our argument is invalid.

Our next example involves *pluralism* and *monism*:

- *Pluralism* (there's more than one being): (∃x)(∃y)~x=y For some x and some y: x≠y

- *Monism* (there's exactly one being): (∃x)(y)y=x For some x, every y is identical to x

Here's a proof that pluralism entails the falsity of monism: 0211

There's more than one being.
∴ It's false that there's exactly one being.

* 1 (∃x)(∃y)~x=y **Valid**
[∴ ~(∃x)(y)y=x
* 2 ┌ asm: (∃x)(y)y=x
* 3 │ ∴ (∃y)~a=y {from 1}
4 │ **∴ ~a=b {from 3}**
5 │ ∴ (y)y=c {from 2}
6 │ ∴ a=c {from 5}
7 │ ∴ b=c {from 5}
8 └ **∴ a=b {from 6 and 7}**
9 ∴ ~(∃x)(y)y=x {from 2; 4 contradicts 8}

Lines 1 and 2 have back-to-back quantifiers. We can drop only quantifiers that are initial and hence outermost; so we drop quantifiers one at a time, starting from the outside. After dropping quantifiers, we substitute equals to get line 8: our "b=c" line lets us take "a=c" and replace "c" with "b," getting "a=b."

We didn't bother to derive "c=c" from "(y)y=c" in line 5. From now on, it'll often be too tedious to drop universal quantifiers using *every* old constant. So we'll just derive instances likely to be useful for our proof or refutation.

Our substitute-equals rule seems to hold universally in arguments about matter or math. But it can fail with mental phenomena. Consider this argument (where "Bx" stands for "Jones believes that x is on the penny"):[1]

Jones believes that Lincoln is on the penny.
Lincoln is the first Republican president.
∴ Jones believes that the first Republican president is on the penny.

Bl
l=r
∴ Br

[1] We could make the same point using relational logic ("Bjl, l=r ∴ Bjr" – where "Bxy" means "x believes that y is on the penny") or Chapter 13's belief logic ("j:Pl, l=r ∴ j:Pr"). Our belief logic explicitly restricts the use of the substitute-equals rule with belief formulas (§13.2).

If Jones is unaware that Lincoln was the first Republican president, the premises could be true while the conclusion is false. So the argument is invalid. But yet we can derive the conclusion from the premises using our substitute-equals rule. So something is wrong here.

To avoid this problem, we'll disallow translating into quantificational logic any predicates or relations that violate the substitute-equals rule. So we won't let "Bx" stand for "Jones believes that x is on the penny." Statements about beliefs and other mental phenomena often violate this rule; so we have to be careful translating such statements into quantificational logic.

So the mental seems to follow different logical patterns from the physical. Does this refute the materialist project of reducing the mental to the physical? Philosophers dispute this question. 0212

9.2a Exercise: LogiCola IDC

Say whether each is valid (and give a proof) or invalid (and give a refutation).

a=b
∴ b=a

1 a=b **Valid**
[∴ b=a
2 ┌ asm: ~b=a
3 │ **∴ ~b=b {from 1 and 2}**
4 └ **∴ b=b {self-identity, to contradict 3}**
5 ∴ b=a {from 2; 3 contradicts 4}

1. Fa
∴ ~(∃x)(Fx • ~x=a)

2. (a=b ⊃ ~(∃x)Fx)
∴ (Fa ⊃ ~Fb)

3. a=b
b=c
∴ a=c

4. ~a=b
c=b
∴ ~a=c

5. ∼a=b
∼c=b
∴ a=c

6. a=b
∴ (Fa ≡ Fb)

7. a=b
(x)(Fx ⊃ Gx)
∼Ga
∴ ∼Fb

8. Fa
∴ (x)(x=a ⊃ Fx)

9. ∴ (∃x)(y)y=x

10. ∴ (∃x)(∃y)∼y=x

9.2b Exercise: LogiCola IDC

First appraise intuitively. Then translate into logic and say whether valid (and give a proof) or invalid (and give a refutation). You'll have to figure out what letters to use; be careful about deciding between small and capital letters.

1. Keith is my only nephew.
My only nephew knows more about BASIC than I do.
Keith is a ten-year-old.
∴ Some ten-year-olds know more about BASIC than I do.

2. Some are logicians.
Some aren't logicians.
∴ There's more than one being.

3. This chemical process is publicly observable.
This pain isn't publicly observable.
∴ This pain isn't identical to this chemical process. [This attacks the identity theory of the mind, which identifies mental events with chemical processes.]

4. The person who left a lighter is the murderer.
The person who left a lighter is a smoker.
No smokers are backpackers.
∴ The murderer isn't a backpacker.

5. The murderer isn't a backpacker.
You aren't a backpacker.
∴ You're the murderer. 0213

6. If Speedy Jones looks back to the quarterback just before the hike, then Speedy Jones is the primary receiver.
The primary receiver is the receiver you should try to cover.
∴ If Speedy Jones looks back to the quarterback just before the hike, then Speedy Jones is the receiver you should try to cover.

7. Judy isn't the world's best cook.
The world's best cook lives in Detroit.
∴ Judy doesn't live in Detroit.

8. Patricia lives in North Dakota.
Blondie lives in North Dakota.
∴ At least two people live in North Dakota.

9. Your grade is the average of your tests.
The average of your tests is B.
∴ Your grade is B.

10. Either you knew where the money was, or the thief knew where it was.
You didn't know where the money was.
∴ You aren't the thief.

11. The man of Suzy's dreams is either rich or handsome.
You aren't rich.
∴ If you're handsome, then you're the man of Suzy's dreams.

12. If someone confesses, then someone goes to jail.
I confess.
I don't go to jail.
∴ Someone besides me goes to jail.

13. David stole money.
The nastiest person at the party stole money.
David isn't the nastiest person at the party.
∴ At least two people stole money. [See problem 4 of §2.3b.]

14. No one besides Carol and the detective had a key.
Someone who had a key stole money.
∴ Either Carol or the detective stole money.

15. Exactly one person lives in North Dakota.
Paul lives in North Dakota.
Paul is a farmer.
∴ Everyone who lives in North Dakota is a farmer.

16. The wildcard team with the best record goes to the playoffs.
Cleveland isn't the wildcard team with the best record.
∴ Cleveland doesn't go to the playoffs.

17. If the thief is intelligent, then you aren't the thief.
∴ You aren't intelligent. 0214

18. You aren't intelligent.
∴ If the thief is intelligent, then you aren't the thief.

9.3 Easier relations

Our final rule for forming quantificational wffs introduces relations:

> The result of writing a capital letter and then two or more small letters is a wff.

Here are two examples:

Lrj = Romeo loves Juliet
Gxyz = x gave y to z

Translating relational sentences into logic can be difficult. We have to study examples and catch patterns; paraphrasing into Loglish is helpful too. We'll start with easier translations and put off multiple-quantifier relations until the next section.

Here are further examples without quantifiers:

Juliet loves Romeo = Ljr
Juliet loves herself = Ljj
Juliet loves Romeo but not Paris = (Ljr • ∼Ljp)

Here are some easy examples with quantifiers:

Everyone loves him/herself = (x)Lxx
Someone loves him/herself = (∃x)Lxx

No one loves him/herself = ~(∃x)Lxx

Normally put quantifiers *before* relations:

Someone (everyone, no one) loves Romeo = For some (all, no) x, x loves Romeo.
Romeo loves someone (everyone, no one) = For some (all, no) x, Romeo loves x.

In the second box, English puts the quantifier last – but logic puts it first. Here are fuller translations: 0215

Someone loves Romeo = (∃x)Lxr For some x, x loves Romeo Everyone loves Romeo = (x)Lxr For all x, x loves Romeo No one loves Romeo = ~(∃x)Lxr It's false that, for some x, x loves Romeo
Romeo loves someone = (∃x)Lrx For some x, Romeo loves x Romeo loves everyone = (x)Lrx For all x, Romeo loves x Romeo loves no one = ~(∃x)Lrx It's false that, for some x, Romeo loves x

These examples are more complicated:

Some Montague loves Juliet
= $(\exists x)(Mx \bullet Lxj)$
For some x, x is a Montague and x loves Juliet

All Montagues love Juliet
= $(x)(Mx \supset Lxj)$
For all x, if x is a Montague then x loves Juliet

Romeo loves some Capulet
= $(\exists x)(Cx \bullet Lrx)$
For some x, x is a Capulet and Romeo loves x

Romeo loves all Capulets
= $(x)(Cx \supset Lrx)$
For all x, if x is a Capulet then Romeo loves x

Here are further examples:

Some Montague besides Romeo loves Juliet
= $(\exists x)((Mx \bullet {\sim}x{=}r) \bullet Lxj)$
For some x, x is a Montague and x ≠ Romeo and x loves Juliet

Romeo loves all Capulets besides Juliet
= $(x)((Cx \bullet {\sim}x{=}j) \supset Lrx)$
For all x, if x is a Capulet and x ≠ Juliet then Romeo loves x

Romeo loves all Capulets who love themselves
= $(x)((Cx \bullet Lxx) \supset Lrx)$
For all x, if x is a Capulet and x loves x then Romeo loves x

Finally, these examples have two different relations:

All who know Juliet love Juliet
= $(x)(Kxj \supset Lxj)$
For all x, if x knows Juliet then x loves Juliet

All who know themselves love themselves
= $(x)(Kxx \supset Lxx)$
For all x, if x knows x then x loves x

Try to master these before starting into the harder relational translations.

9.3a Exercise: LogiCola H (RM & RT)

Using these equivalences, translate these English sentences into wffs. 0216

Lxy = x loves y
Cxy = x caused y
Gxy = x is greater than y

Ix = x is Italian
Rx = x is Russian
Ex = x is evil

t = Tony
o = Olga
g = God

God caused nothing that is evil.
∼(∃x)(Ex • Cgx)

1. Tony loves Olga and Olga loves Tony.
2. Not every Russian loves Olga.
3. Tony loves everyone who is Russian.
4. Olga loves someone who isn't Italian.
5. Everyone loves Olga but not everyone is loved by Olga.
6. All Italians love themselves.
7. Olga loves every Italian besides Tony.
8. Tony loves everyone who loves Olga.
9. No Russian besides Olga loves Tony.
10. Olga loves all who love themselves.
11. Tony loves no Russians who love themselves.
12. Olga is loved.
13. God caused everything besides himself.
14. Nothing caused God.
15. Everything that God caused is loved by God.
16. Nothing caused itself.
17. God loves himself.
18. If God did not cause himself, then there is something that God did not cause.
19. Nothing is greater than God.
20. God is greater than anything that he caused.

9.4 Harder relations

Now we get into multiple-quantifier translations. Here's a simple example:

Someone loves someone
= (∃x)(∃y)Lxy
For some x and for some y, x loves y

This could be true because some love themselves ("(∃x)Lxx") or because some love another ("(∃x)(∃y)(∼x=y • Lxy)"). Here are more examples:

Everyone loves everyone
= (x)(y)Lxy
For all x and for all y, x loves y

Some Montague hates some Capulet
= (∃x)(∃y)((Mx • Cy) • Hxy)
For some x and for some y, x is a Montague and y is a Capulet and x hates y

Every Montague hates every Capulet
= (x)(y)((Mx • Cy) ⊃ Hxy)
For all x and for all y, if x is a Montague and y is a Capulet then x hates y 0217

Study carefully this next pair – which differs only in the quantifier order:

Everyone loves someone. For all x there's some y, such that x loves y. (x)(∃y)Lxy
There's someone whom everyone loves. There's some y such that, for all x, x loves y. (∃y)(x)Lxy

In the first case, we might love *different* people. In the second, we love the *same* person; perhaps we all love God. Notice the difference in the contrasting pairs:

Everyone loves someone ≠ There's someone whom everyone loves

Everyone lives in some house ≠ There's some house where everyone lives

Everyone has some job ≠ There's some job that everyone has

Everyone makes some error ≠ There's some error that everyone makes

The sentences on the right make a stronger claim: some-every entails every-some, but not the other way around.

Back-to-back quantifiers of the same type can be switched: "(x)(y)" = "(y)(x)" and "(∃x)(∃y)" = "(∃y)(∃x)." But the order matters if the quantifiers are of different types: "(∃x)(y)" is stronger than "(y)(∃x)." It doesn't matter what variable letters we use, so long as the reference pattern is the same. So in "(x)(∃y)Lxy" we could use other variables in place of "x" and "y" – as long our wff consists in a universal and then an existential (using different variables), "L," the variable used in the universal, and finally the variable used in the existential.

Here's a difficult every-some translation, which we'll do step by step:

Every Capulet loves some Montague
For all x, if x is a Capulet then x loves some Montague
(x)(Cx ⊃ x loves some Montague)
(x)(Cx ⊃ for some y, y is a Montague and x loves y)
(x)(Cx ⊃ (∃y)(My • Lxy))

Until you master these, go by "baby steps" from English to Loglish to symbols. First go from "*Every Capulet* loves some Montague" to "For all x, *if* x is a Capulet *then* x loves some Montague"; so the formula after "(x)" is an IF-THEN. Later go from "x loves *some Montague*" to "for some y, y is a Montague *and* x loves y"; this part is an AND. So "*every* Capulet" gives "*if* Capulet *then* ...," and "*some* Montague" gives "some are Montague *and* ..."

As usual, we could switch conjuncts ("My" and "Lxy"). We also could put the existential further out, so the wff starts "(x)(∃y)." But the order of the quantifiers has to follow the English – so if "every" comes before "some" then "(x)" has to come before "(∃y)." This example is difficult; study it carefully.

Here are analogous but easier every-some translations: 0218

Every Capulet loves someone
For all x, if x is a Capulet then x loves someone
(x)(Cx ⊃ x loves someone)
(x)(Cx ⊃ for some y, x loves y)
(x)(Cx ⊃ (∃y)Lxy)

Everyone loves some Montague
For all x, x loves some Montague
(x) x loves some Montague
(x) for some y, y is a Montague and x loves y
(x)(∃y)(My • Lxy)

The first uses IF-THEN, because "*Every Capulet* loves someone" goes into "For all x, *if* x is a Capulet *then* x loves someone." The second uses AND, because "x loves *some Montague*" goes into "for some y, y is a Montague *and* x loves y."

Here's a difficult some-every translation:

Some Capulet loves every Montague
For some x, x is a Capulet and x loves every Montague
(∃x)(Cx • x loves every Montague)
(∃x)(Cx • for all y, if y is a Montague then x loves y)
(∃x)(Cx • (y)(My ⊃ Lxy))

"*Some Capulet* loves every Montague" becomes "For some x, x is a Capulet *and* x loves every Montague"; "(∃x)" is followed by AND. Then "x loves *every Montague*" becomes "for all y, *if* y is a Montague *then* x loves y"; "(x)" is followed by IF-THEN. As before, we could switch conjuncts "Cx" and "(y)(My ⊃ Lxy)." And we could start the wff with "(∃x)(y)"; here "(∃x)" has to come before "(y)," since "some" comes before "every" in the English.

Here are analogous but easier some-every translations:

Some Capulet loves everyone
For some x, x is a Capulet and x loves everyone
(∃x)(Cx • x loves everyone)
(∃x)(Cx • for all y, x loves y)
(∃x)(Cx • (y)Lxy)

Someone loves every Montague
For some x, x loves every Montague
(∃x) x loves every Montague
(∃x) for all y, if y is a Montague then x loves y
(∃x)(y)(My ⊃ Lxy)

The first uses AND, because "*Some Capulet* loves everyone" becomes "For some x, x is a Capulet *and* x loves everyone." The second uses IF-THEN, because "x loves *every Montague*" becomes "for all y, *if* y is a Montague *then* x loves y." 0219

Since these translations are difficult, you might want to reread a couple of times from the beginning of this section to here, until you get it.

Here are some miscellaneous translations:

There's an unloved lover
= $(\exists x)(\sim(\exists y)Lyx \bullet (\exists y)Lxy)$
For some x, x is unloved (no one loves x) and x is a lover (x loves someone)

Everyone loves a lover
= $(x)((\exists y)Lxy \supset (y)Lyx)$
For all x, if x is a lover (x loves someone) then everyone loves x

Romeo loves all and only those who don't love themselves
= $(x)(Lrx \equiv \sim Lxx)$
For all x, Romeo loves x if and only if x doesn't love x

All who know any person love that person
= $(x)(y)(Kxy \supset Lxy)$
For all x and all y, if x knows y then x loves y

Relations have properties like reflexivity, symmetry, and transitivity:

"Having the same age as" is *reflexive*
= $(x)Axx$
Everything has the same age as itself

"Being taller than" is *irreflexive*
= $\sim(\exists x)Txx$
Nothing is taller than itself

"Being a relative of" is *symmetrical*
= $(x)(y)(Rxy \supset Ryx)$
In all cases, if x is a relative of y, then y is a relative of x

"Being a father of" is *asymmetrical*
= $(x)(y)(Fxy \supset \sim Fyx)$
In all cases, if x is a father of y then y isn't a father of x

"Being taller than" is *transitive*
= $(x)(y)(z)((Txy \bullet Tyz) \supset Txz)$
In all cases, if x is taller than y and y is taller than z, then x is taller than z

"Being a father of" is *intransitive*
= $(x)(y)(z)((Fxy \bullet Fyz) \supset \sim Fxz)$
In all cases, if x is a father of y and y is a father of z, then x isn't a father of z

Love fits none of these six categories. Love is neither reflexive nor irreflexive: sometimes people love themselves and sometimes they don't. Love is neither symmetrical nor asymmetrical: if x loves y, then sometimes y loves x in return and sometimes not. And love is neither transitive nor intransitive: if x loves y and y loves z, then sometimes x loves z and sometimes not.

9.4a Exercise: LogiCola H (RM & RT)

Using these equivalences, translate these English sentences into wffs. 0220

Lxy = x loves y
Cxy = x caused y
Gxy = x is greater than y

Ix = x is Italian
Rx = x is Russian
Ex = x is evil

t = Tony
o = Olga

Every Russian loves everyone.
(x)(Rx ⊃ (y)Lxy) *or* (x)(y)(Rx ⊃ Lxy)

1. Everyone loves every Russian.
2. Some Russians love someone.
3. Someone loves some Russians.
4. Some Russians love every Italian.
5. Every Russian loves some Italian.
6. There is some Italian that every Russian loves.
7. Everyone loves everyone else.
8. Every Italian loves every other Italian.
9. Some Italians love no one.
10. No Italians love everyone.
11. No one loves all Italians.
12. Someone loves no Italians.
13. No Russians love all Italians.
14. If everyone loves Olga, then there is some Russian that everyone loves.
15. If Tony loves everyone, then there is some Italian who loves everyone.

16. It is not always true that if a first thing caused a second, then the first is greater than the second.

17. In all cases, if a first thing is greater than a second, then the second isn't greater than the first.

18. Everything is greater than something.

19. There's something than which nothing is greater.

20. Everything is caused by something.

21. There's something that caused everything.

22. Something evil caused all evil things.

23. In all cases, if a first thing caused a second and the second caused a third, then the first caused the third.

24. There's a first cause (there's some x that caused something but nothing caused x).

25. Anyone who caused anything loves that thing.

9.5 Relational proofs

In relational proofs, as before, we'll reverse squiggles, then drop existentials, and lastly drop universals. Drop only *initial* (outermost) quantifiers. So with back-to-back quantifiers "(x)(y)" (in line 3 below), drop "(x)" first and then "(y)": 0221

Paris loves Juliet.
Juliet doesn't love Paris.
∴ It's not always true that if a first person loves a second then the second loves the first.

```
   1  Lpj    Valid
   2  ~Ljp
   [ ∴ ~(x)(y)(Lxy ⊃ Lyx)
   3  ┌ asm: (x)(y)(Lxy ⊃ Lyx)
   4  │ ∴ (y)(Lpy ⊃ Lyp)    {from 3}
 * 5  │ ∴ (Lpj ⊃ Ljp)    {from 4}
   6  └ ∴ Ljp    {from 1 and 5}
   7 ∴ ~(x)(y)(Lxy ⊃ Lyx)    {from 3; 2 contradicts 6}
```

Our older proof strategy would have us drop each initial universal quantifier twice, once using "p" and once using "j." But now this would be tedious; so henceforth we'll derive only what will be useful for our proof or refutation.

Here's another relational proof:

There's someone that everyone loves.
∴ Everyone loves someone.

```
* 1   (∃y)(x)Lxy    Valid
   [ ∴ (x)(∃y)Lxy
* 2 ┌ asm: ~(x)(∃y)Lxy
* 3 │ ∴ (∃x)~(∃y)Lxy    {from 2}
* 4 │ ∴ ~(∃y)Lay    {from 3}
  5 │ ∴ (y)~Lay    {from 4}
  6 │ ∴ (x)Lxb    {from 1}
  7 │ ∴ Lab    {from 6}
  8 └ ∴ ~Lab    {from 5}
  9 ∴ (x)(∃y)Lxy    {from 2; 7 contradicts 8}
```

This is valid intuitively: if there's one specific person (God, for example) that everyone loves, then everyone loves at least one person.

For quantificational arguments without relations and identity:

1. there are mechanical strategies (like that sketched in §8.2) that always give a proof or refutation in a finite number of lines; and

2. a refutation at most needs 2n entities (where n is the number of distinct predicates in the argument).

Neither holds for relational arguments. Against 1, no possible mechanical strategy will always give a proof or refutation of a relational argument. This result is called *Church's theorem*, after Alonzo Church. So working out relational arguments sometimes requires ingenuity and not just mechanical methods; the problem with our proof strategy is that it can lead into endless loops.[1] Against 2, refuting invalid relational arguments sometimes requires a possible world with an infinite number of entities. 0222

Instructions lead into an **endless loop** if they command the same sequence of actions over and over, endlessly. I've written computer programs with endless loops by mistake. I put an endless loop into the Index for fun:

Endless loop; see loop, endless
Loop, endless; see endless loop

[1] The companion LogiCola computer program follows mechanical rules (algorithms) to construct proofs. Left to itself, it would go into an endless loop for some invalid relational arguments. But LogiCola is told beforehand which arguments go into an endless loop and which refutations to then give, so it can stop the loop at a reasonable point.

Our quantificational proof strategy can lead into such a loop. If you see this coming, quit the strategy and improvise your own refutation.

Wffs that begin with a universal/existential quantifier combination, like "(x)(∃y)," often lead into endless loops. Here's an example:[1]

Everyone loves someone.
∴ There's someone that everyone loves.

(x)(∃y)Lxy **Invalid**
∴ (∃y)(x)Lxy

The premise by itself leads into an endless loop:

Everyone loves someone.
∴ a loves someone.
∴ a loves b.
∴ b loves someone.
∴ b loves c.
∴ c loves someone.
∴ c loves d.
… and so on endlessly …

(x)(∃y)Lxy
∴ (∃y)Lay
∴ Lab
∴ (∃y)Lby
∴ Lbc
∴ (∃y)Lcy
∴ Lcd
… and so on endlessly …

This argument is invalid, but we have to improvise to get the refutation; we can't wait until our proof strategy ends (since it never will) and then use the simple formulas to construct a refutation. Instead, we have to think out the refutation by ourselves. While there's no strategy that always works, I suggest that you:

- break out of the loop before introducing your third constant (often it suffices

[1] This example is like arguing "Everyone lives in some house, so there must be some (one) house that everyone lives in." Some great minds have committed this *quantifier-shift fallacy*. Aristotle argued, "Every agent acts for an end, so there must be some (one) end for which every agent acts." St Thomas Aquinas argued, "If everything at some time fails to exist, then there must be some (one) time at which everything fails to exist." And John Locke argued, "Everything is caused by something, so there must be some (one) thing that caused every-thing."

to use two beings, a and b; don't multiply entities unnecessarily),

- begin your refutation with values you already have (maybe you already have "Lab" and "Laa"), and
- add other wffs to make premises true and conclusion false (maybe try adding "Lba" or "∼Lba," and then "Lbb" or "∼Lbb," until your refutation works).

Fiddle with the values until you find a refutation that works. 0223

Consider our example again:

Everyone loves someone.
∴ There's someone that everyone loves.

(x)(∃y)Lxy **Invalid**
∴ (∃y)(x)Lxy

If we stop the attempted proof before introducing our third constant, we may get either of these as the *beginning* of our refutation:

a, b	∼Lab

a, b	Lab

We need to add more formulas to make the premise true and conclusion false. With the first box, we need EVERYONE to love someone. Since a doesn't love b, we need to have a love a. So we add this:

a, b	Laa, ∼Lab

We also need b to love someone (so we need Lbb or Lba) – but without there being some one person that everyone loves (which excludes Lba, so we have to add Lbb and ∼Lba). So with ingenuity we construct this possible world, with beings a and b, that makes the premise true and conclusion false:

a, b	Laa, ∼Lab Lbb, ∼Lba

In this egoistic world, all love themselves but not others. This makes "Everyone loves someone" true but "There's someone that everyone loves" false (since not everyone loves a and not everyone loves b). This refutation works too:

a, b	Lab, ∼Laa Lba, ∼Lbb

In this altruistic world, all love others but not themselves. This makes "Everyone loves someone" true but "There's someone that everyone loves" false (since not everyone loves a and not everyone loves b).

Refuting relational arguments sometimes requires a universe with an infinite number of entities. Here's an example:

In all cases, if x is greater than y and y is greater than z then x is greater than z.
In all cases, if x is greater than y then y isn't greater than x.
b is greater than a.
∴ There's something than which nothing is greater.

(x)(y)(z)((Gxy • Gyz) ⊃ Gxz) **Invalid**
(x)(y)(Gxy ⊃ ∼Gyx)
Gba
∴ (∃x)∼(∃y)Gyx 0224

We can imagine a world with an infinity of beings – in which each being is surpassed in greatness by another. Let's take the natural numbers (0, 1, 2, ...) as the universe of discourse. Let "a" refer to 0 and "b" refer to 1 and "Gxy" mean "x > y." On this interpretation, the premises are all true. But the conclusion, which says "There's a number than which no number is greater," is false. This shows that the form is invalid.

So relational arguments raise problems about infinity (endless loops and infinite worlds) that other kinds of argument we've studied don't raise.

9.5a Exercise: LogiCola I (RC & BC)

Say whether each is valid (and give a proof) or invalid (and give a refutation).

(∃x)(∃y)Lxy
∴ (∃y)(∃x)Lxy

* 1   (∃x)(∃y)Lxy   **Valid**
   [∴ (∃y)(∃x)Lxy
* 2 ┌ asm: ∼(∃y)(∃x)Lxy
   3 │ ∴ (y)∼(∃x)Lxy   {from 2}
* 4 │ ∴ (∃y)Lay   {from 1}
   5 │ **∴ Lab   {from 4}**
* 6 │ ∴ ∼(∃x)Lxb   {from 3}
   7 │ ∴ (x)∼Lxb   {from 6}
   8 └ **∴ ∼Lab   {from 7}**
   9 ∴ (∃y)(∃x)Lxy   {from 2; 5 contradicts 8}

1. (x)Lxa
∴ (x)Lax

2. (∃x)(y)Lxy
∴ (∃x)Lxa

3. (x)(y)(Lxy ⊃ x=y)
∴ (x)Lxx

4. (x)(∃y)Lxy
∴ Laa

5. (x)(y)Lxy
∴ (x)(y)((Fx • Gy) ⊃ Lxy)

6. (x)(y)(Uxy ⊃ Lxy)
(x)(∃y)Uxy
∴ (x)(∃y)Lxy

7. (x)Lxx
∴ (∃x)(y)Lxy

8. (x)Gaxb
∴ (∃x)(∃y)Gxcy

9. (x)(y)Lxy
∴ (∃x)Lax

10. Lab
Lbc
∴ (∃x)(Lax • Lxc)

11. (x)Lxx
∴ (x)(y)(Lxy ⊃ x=y)

12. (∃x)Lxa
~Laa
∴ (∃x)(~a=x • Lxa)

13. (x)(y)(z)((Lxy • Lyz) ⊃ Lxz)
(x)(y)(Kxy ⊃ Lyx)
∴ (x)Lxx

14. (x)Lxa
(x)(Lax ⊃ x=b)
∴ (x)Lxb

15. (x)(y)(Lxy ⊃ (Fx • ~Fy))
∴ (x)(y)(Lxy ⊃ ~Lyx) 0225

9.5b Exercise: LogiCola I (RC & BC)

First appraise intuitively. Then translate into logic and say whether valid (and give a proof) or invalid (and give a refutation).

1. Juliet loves everyone.
∴ Someone loves you. [Use Lxy, j, and u.]

2. Nothing caused itself.
∴ There's nothing that caused everything. [Use Cxy.]

3. Alice is older than Betty.
∴ Betty isn't older than Alice. [Use Oxy, a, and b. What implicit premise would make this valid?]

4. There's something that everything depends on.
∴ Everything depends on something. [Dxy]

5. Everything depends on something.
∴ There's something that everything depends on. [Dxy]

6. Paris loves all females.
No females love Paris.
Juliet is female.
∴ Paris loves someone who doesn't love him.   [Lxy, p, Fx, j]

7. In all cases, if a first thing caused a second, then the first exists before the second.
Nothing exists before it exists.
∴ Nothing caused itself.   [Use Cxy and Bxy (for "x exists before y exists").]

8. Everyone hates my enemy.
My enemy hates no one besides me.
∴ My enemy is me.   [Hxy, e, m]

9. Not everyone loves everyone.
∴ Not everyone loves you.   [Lxy, u]

10. There's someone that everyone loves.
∴ Some love themselves.

11. Andy shaves all and only those who don't shave themselves.
∴ It is raining.   [Sxy, a, R]

12. No one hates themselves.
I hate all logicians.
∴ I am not a logician.   [Hxy, i, Lx]

13. Juliet loves everyone besides herself.
Juliet is Italian.
Romeo is my logic teacher.
My logic teacher isn't Italian.
∴ Juliet loves Romeo.   [j, Lxy, Ix, r, m] 0226

14. Romeo loves either Lisa or Colleen.
Romeo doesn't love anyone who isn't Italian.
Colleen isn't Italian.
∴ Romeo loves Lisa.   [Lxy, r, l, c]

15. Everyone loves all lovers.
Romeo loves Juliet.
∴ I love you.   [Use Lxy, r, j, i, and u. This one is difficult.]

16. Everyone loves someone.
∴ Some love themselves.

17. Nothing caused itself.
This chemical brain process caused this pain.
∴ This chemical brain process isn't identical to this pain. [Cxy, b, p]

18. For every positive contingent truth, something explains why it's true.
The existence of the world is a positive contingent truth.
If something explains the existence of the world, then some necessary being explains the existence of the world.
∴ Some necessary being explains the existence of the world. [Use Cx, Exy, e, and Nx. This argument for the existence of God is from Richard Taylor.]

19. That girl is Miss Novak.
∴ If you don't like Miss Novak, then you don't like that girl. [Use t, m, u, and Lxy; from the movie, *The Little Shop around the Corner*: "If you don't like Miss Novak, I can tell you right now that you won't like that girl. Why? Because it is Miss Novak."]

20. Everyone who is wholly good prevents every evil that he can prevent.
Everyone who is omnipotent can prevent every evil.
If someone prevents every evil, then there's no evil.
There's evil.
∴ Either God isn't omnipotent, or God isn't wholly good. [Use Gx, Ex, Cxy (for "x can prevent y"), Pxy (for "x prevents y"), Ox, and g; from J. L. Mackie.]

21. Your friend is wholly good.
Your knee pain is evil.
Your friend can prevent your knee pain.
Your friend doesn't prevent your knee pain (since he could prevent it only by amputating your leg – which would bring about a worse situation).
∴ "Everyone who is wholly good prevents every evil that he can prevent" is false. [Use f, Gx, k, Ex, Cxy, and Pxy. Alvin Plantinga thus attacked premise 1 of the previous argument; he proposed instead roughly this: "Everyone who is wholly good prevents every evil that he knows about if he can do so without thereby eliminating a greater good or bringing about a greater evil."] 0227

22. For everything contingent, there's some time at which it fails to exist.
∴ If everything is contingent, then there's some time at which everything fails to exist. [Use Cx for "x is contingent"; Ext for "x exists at time t"; t for a time variable; and t´, t´´, t´´´, … for time constants. This is a critical step in St Thomas Aquinas's third argument for the existence of God.]

23. If everything is contingent, then there's some time at which everything fails to exist.
If there's some time at which everything fails to exist, then there's nothing in existence now.
There's something in existence now.
Everything that isn't contingent is necessary.
∴ There's a necessary being. [Besides the letters for the previous argument, use Nx for "x is necessary" and n for "now." This continues Aquinas's argument; here premise 1 is from the previous argument.]

24. [Gottlob Frege tried to systematize mathematics. One of his axioms said that *every sentence with a free variable*[1] *determines a set*; so "x is blue" determines a set containing all and only blue things. While this seems sensible, Bertrand Russell showed that this entails that "x doesn't contain x" determines a set y containing all and only those things that don't contain themselves – and this leads to the self-contradiction "y contains y if and only if y doesn't contain y." The foundations of mathematics haven't been the same since "Russell's paradox."]
If every sentence with a free variable determines a set, then there's a set y such that, for all x, y contains x if and only if x doesn't contain x.
∴ Not every sentence with a free variable determines a set. [Use D for "Every sentence with a free variable determines a set," Sx for "x is a set," and Cyx for "y contains x." See §16.4.]

25. All dogs are animals.
∴ All heads of dogs are heads of animals. [Use Dx, Ax, and Hxy (for "x is a head of y"). Translate "x is a head of a dog" as "for some y, y is a dog and x is a head of y." Augustus De Morgan in the 19th century claimed that this was a valid argument that traditional logic couldn't validate.]

9.6 Definite descriptions

Definite descriptions, phrases of the form "the so and so," are used to pick out a definite (single) person or thing. Consider how we've been symbolizing these two English sentences:

Socrates is bald = Bs
The king of France is bald = Bk 0228

The first sentence has a proper name ("Socrates") while the second has a definite description ("the king of France"); both seem to ascribe a *property* (baldness) to a particular *object* or entity. Bertrand Russell argued that this

[1] An instance of a variable is "free" in a wff if it doesn't occur as part of a wff that begins with a quantifier using that variable; each instance of "x" is free in "Fx" but not in "(x)Fx."

object-property analysis is misleading. Definite descriptions (like "the king of France") should instead be analyzed using a complex of predicates and quantifiers:

The king of France is bald
= $(\exists x)((Kx \bullet \sim(\exists y)(\sim y{=}x \bullet Ky)) \bullet Bx)$
There's exactly one king of France, and he's bald
For some x, x is king of France and there's no y such that: y≠x and y is king of France and x is bald

Russell saw his analysis as having two advantages.

First, "The king of France is bald" might be false for any of three reasons:

1. There's no king of France;
2. there's more than one king of France; or
3. there's exactly one king of France, and he has hair on his head.

In fact, "The king of France is bald" is false for reason 1: France has no king. This fits Russell's analysis. By contrast, the object-property analysis suggests that if "The king of France is bald" is false, then "The king of France isn't bald" would have to be true – and so the king of France would have to have hair! So Russell's analysis expresses better the logical complexity of definite descriptions.

Second, the object-property analysis of definite descriptions can easily lead us into metaphysical errors, like positing existing things that aren't real. The philosopher Alexius Meinong argued roughly as follows (and Russell at first accepted this argument):

"The round square does not exist" is a true statement about the round square.
If there's a true statement about something, then that something has to exist.
∴ The round square exists.
But the round square isn't a real thing.
∴ Some things that exist aren't real things.

Russell later saw the belief in existing non-real things as foolish. Appealing to his theory of descriptions, he criticized Meinong's argument as resting on a naïve object-property analysis of this statement:

"The round square does not exist."

If this were a true statement about the round square, as Meinong's first premise asserts, then the round square would have to exist – which the statement denies. Instead, the statement just denies that there's exactly one being that's both round and square. So Russell's analysis keeps us from having to accept existing things that aren't real. 0229

9.7 Copi proofs

We earlier discussed traditional Copi proofs for propositional logic and basic quantificational logic (§§7.5 and 8.6). Copi proofs can also be used for identity and relations. Copi uses our SI (Self-identity) and SE (Substitute Equals) rules, and adds a SS (Switch Sides) replacement rule: "x=y = y=x" (where you can use any variable or constant for "x" and for "y").

9.7a and 9.7b Exercise: LogiCola IDO

Do Copi proofs for problems in §9.2a (just the valid ones, namely 3, 4, 6, 7, and 8) and §9.2b (just the valid ones, namely 1, 2, 3, 4, 6, 9, 10, 12, 13, 14, 15, and 18). These are identity arguments.

9.7c and 9.7d Exercise: LogiCola IRO and IBO

Do Copi proofs for problems in §9.5a (just the valid ones, namely 2, 5, 6, 8, 9, 10, 12, 14, and 15) and §9.5b (just the valid ones, namely 1, 2, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 23, 24, and 25). These are relational arguments.

10 Basic Modal Logic

Modal logic studies arguments whose validity depends on "possible," "necessary," and similar notions. This chapter covers the basics and the next gets into further modal systems.

10.1 Translations

To help us evaluate arguments, we'll construct a modal language. This includes propositional logic's vocabulary, wffs, inference rules, and proofs. It adds symbols for *modal operators*: "◇" and "☐" (diamond and box):

◇A = It's possible that A. A is true in some possible world.
A = It's true that A. A is true in the actual world.
☐A = It's necessary that A. A is true in all possible worlds.

"Possible" is weaker than "true," while "necessary" is stronger than "true." "A is necessary" claims that A *has* to be true – it *couldn't* have been false.

"Possible" here means *logically possible* (*not self-contradictory*). "I run a mile in two minutes" may be physically impossible; but it's logically possible (the idea contains no self-contradiction). Likewise, "necessary" means *logically necessary* (*self-contradictory to deny*). "2+2 = 4" and "All bachelors are unmarried" are examples of **necessary truths**; such truths are based on logic, the meaning of concepts, or necessary connections between properties.

We can rephrase "possible" as *true in some possible world* – and "necessary" as *true in all possible worlds*. A **possible world** is a consistent description of how things might have been or might in fact be. Picture a possible world as a *consistent story* (or novel). The story is *consistent*, in that its statements don't entail self-contradictions; it describes a set of possible situations that are all possible together. The story may or may not be true. The **actual world** is the story that's true – the description of how things in fact are.

As before, grammatical formulas are *wffs* (*well-formed formulas*). Wffs now are strings we can construct using the propositional rules plus this new rule:

The result of writing "◇" or "☐," and then a wff, is a wff.
0231

Don't use parentheses with "◇A" and "☐A"; these forms are incorrect: "◇(A)," "(◇A)," "☐(A)," "(☐A)." Parentheses here would serve no purpose.

We'll focus now on how to translate English sentences into modal logic. Here are some simpler examples:

A is possible (consistent, could be true)
= ◇A

A is necessary (must be true, has to be true)
= ☐A

A is impossible (self-contradictory)
= ~◇A = A couldn't be true
= ☐~A = A has to be false

An impossible statement (like "2 ≠ 2") is one that's false in every possible world.

These examples are more complicated:

A is consistent (compatible) with B
= ◇(A • B)
It's possible that A and B are both true

A entails B
= ☐(A ⊃ B)
It's necessary that if A then B

"Entails" makes a stronger claim than plain "if-then." Compare these two:

"There's rain" entails "There's precipitation"
= ☐(R ⊃ P)

If it's Saturday, then I don't teach class
= (S ⊃ ~T)

The first is logically necessary; every conceivable situation with rain also has precipitation. The second just happens to be true; we can consistently imagine me teaching on Saturday, even though in fact I never do.

Here are further forms:

A is inconsistent with B
= $\sim\Diamond(A \bullet B)$
It's not possible that A and B are both true

A doesn't entail B
= $\sim\Box(A \supset B)$
It's not necessary that if A then B

A is a contingent statement[1]
= $(\Diamond A \bullet \Diamond\sim A)$
A is possible and not-A is possible

A is a contingent truth
= $(A \bullet \Diamond\sim A)$
A is true but could have been false

Statements are necessary, impossible, or contingent. But truths are only necessary or contingent (since impossible statements are false). 0232

When translating, it's usually good to mimic the English word order:

necessary not	=	$\Box\sim$
not necessary	=	$\sim\Box$

necessary if	=	$\Box($
if necessary	=	$(\Box$

Use a separate box or diamond for each "necessary" or "possible." So "If A is *necessary* and B is *possible*, then C is *possible*" is "$((\Box A \bullet \Diamond B) \supset \Diamond C)$."

This English form is ambiguous between two meanings:

"If you're a bachelor, then you must be unmarried."

[1] English sometimes uses "possible" to mean what we call "contingent" (*true in at least one possible world and false in at least one possible world*). In our sense of "possible" (*true in at least one possible world*), what is necessary is also thereby possible.

Simple necessity $(B \supset \Box U)$ If you're a bachelor, then you're *inherently unmarriable* (in no possible world would anyone marry you). If B, then U (by itself) is necessary.
Conditional necessity $\Box(B \supset U)$ It's necessary that *if* you're a bachelor *then* you're unmarried. It's necessary that if-B-then-U.

Box-inside "$(B \supset \Box U)$" affirms *simple necessity*: given your bachelorhood, "You're unmarried" is inherently necessary; this is insulting and presumably false. **Box-outside** "$\Box(B \supset U)$" affirms *conditional necessity*: what's necessary isn't "You're a bachelor" or "You're unmarried" by itself, but the connection between the two: it's necessary that *if* you're a bachelor (unmarried man) *then* you're unmarried. So our English "*If you're a bachelor, then you must be unmarried*" is ambiguous; its wording suggests simple necessity (which denies your freedom to marry) but it's likely meant as conditional necessity.

Medievals called the box-inside form "necessity of the *consequent*" (the second part is necessary) and the box-outside form "necessity of the *consequence*" (the if-then is necessary). The ambiguity is important; several fallacious philosophical arguments depend on the ambiguity for their plausibility.

It's not ambiguous if you say that the second part is "by itself" or "intrinsically" necessary or impossible – or if you use "entails" or start with "necessary." These forms aren't ambiguous:

If A, then B (by itself) is necessary = $(A \supset \Box B)$
If A, then B is intrinsically necessary = $(A \supset \Box B)$

A entails B = $\Box(A \supset B)$
Necessarily, if A then B = $\Box(A \supset B)$
It's necessary that if A then B = $\Box(A \supset B)$
"If A then B" is a necessary truth = $\Box(A \supset B)$

The ambiguous forms have if-then with a strong modal term (like "necessary," 0233 "must," "impossible," or "can't") in the then-part:[1]

"*If A is true, then it's necessary (must be) that B*" could mean "(A ⊃ ☐B)" or "☐(A ⊃ B)."
"*If A is true, then it's impossible (couldn't be) that B*" could mean "(A ⊃ ☐~B)" or "☐(A ⊃ ~B)."

When you translate an ambiguous English sentence, give both forms. With ambiguous arguments, work out both arguments.

10.1a Exercise: LogiCola J (BM & BT)

Translate these into wffs. Be sure to translate ambiguous forms both ways.

"God exists and evil doesn't exist" entails "There's no matter."
☐((G • ~E) ⊃ ~M)

1. It's necessary that God exists.
2. "There's a God" is self-contradictory.
3. It's not necessary that there's matter.
4. It's necessary that there's no matter.
5. "There's rain" entails "There's precipitation."
6. "There's precipitation" doesn't entail "There's rain."
7. "There's no precipitation" entails "There's no rain."
8. If rain is possible, then precipitation is possible.
9. God exists.
10. If there's rain, then there must be rain.
11. It's not possible that there's evil.
12. It's possible that there's no evil.
13. If you get more points than your opponent, then it's impossible for you to lose.
14. It's necessary that if you see that B is true then B is true.

[1] There's an exception to these boxed rules: if the if-part is a claim about necessity or possibility, then just use the box-inside form. So "If A is necessary then B is necessary" is just "(☐A ⊃ ☐B)" – and "If A is possible then B is impossible" is just "(◇A ⊃ ~◇B)."

15. If B has an all-1 truth table, then B is inherently necessary.
16. Necessarily, if there's a God then there's no evil.
17. If there's a God, then there can't be evil.
18. If there must be matter, then there's evil.
19. Necessarily, if there's a God then "There's evil" (by itself) is self-contradictory.
20. It's necessary that it's heads or tails.
21. Either it's necessary that it's heads or it's necessary that it's tails.
22. "There's rain" is a contingent statement.
23. "There's rain" is a contingent truth.
24. "If there's rain, then there's evil" is a necessary truth.
25. If there's rain, then "There's evil" (by itself) is logically necessary.
26. If there's rain, then it's necessary that there's evil. 0234
27. It's necessary that it's possible that there's matter.
28. "There's a God" isn't a contingent truth.
29. If there's a God, then it must be that there's a God.
30. It's necessary that if there's a God then "There's a God" (by itself) is necessary.

10.2 Proofs

For modal proofs, we need world prefixes and modal inference rules.

A **world prefix** is a string of zero or more instances of "W." So " " (zero instances), "W," "WW," and so on are world prefixes; these represent possible worlds, with the blank world prefix (" ") representing the actual world. A *derived line* is now a line consisting of a world prefix and then "∴" and then a wff. And an *assumption* is now a line consisting of a world prefix and then "asm:" and then a wff. Here are examples of derived lines and assumptions:

∴ A (So A is true in the actual world.)
W ∴ A (So A is true in world W.)
WW ∴ A (So A is true in world WW.)

asm: A (Assume A is true in the actual world.)
W asm: A (Assume A is true in world W.)
WW asm: A (Assume A is true in world WW.)

Derived lines with W's are more common.

We can use S- and I-rules and RAA in modal proofs. Unless otherwise specified, we can use an inference rule only within a given world; so if we have "(A ⊃ B)" and "A" in the same world, then we can infer "B" in this same world. RAA needs additional wording (*italicized below*) for world prefixes:

> RAA: Suppose some pair of not-blocked-off lines *using the same world prefix* has contradictory wffs. Then block off all the lines from the last not-blocked-off assumption on down and infer a line consisting in *this assumption's world prefix followed by* "∴" followed by a contradictory of the assumption.

For RAA, lines with contradictory wffs must have the same world prefix. "W ∴ A" and "WW ∴ ∼A" isn't enough; "A" may be true in one world but false in another. But "WW ∴ A" and "WW ∴ ∼A" is a genuine contradiction. And the derived line must have the same world prefix as the assumption; if "W asm: A" leads to a contradiction in any world, then RAA lets us derive "W ∴ ∼A."

Modal proofs use four new inference rules. The reverse-squiggle rules hold regardless of what pair of contradictory wffs replaces "A" / "∼A"; here "→" means that we can infer whole lines from left to right: 0235

Reverse squiggle RS
∼☐A → ◇∼A
∼◇A → ☐∼A

"Not necessary" entails "possibly false." And "not possible" entails "necessarily false." Use these rules only within the same world. Our rules cover reversing squiggles on longer formulas, if the whole formula begins with "∼☐" or "∼◇":

∼◇∼B
————————
∴ ☐∼∼B

∼☐(C • ∼D)
————————————
∴ ◇∼(C • ∼D)

In the first example, we also could conclude "☐B" (dropping "∼∼"). This next example is illegal in our system, since it fits poorly into our proof strategy, even though it's logically correct:

Don't do this:

(P ⊃ ~☐Q) _
∴ (P ⊃ ◇~Q)

Reverse squiggles whenever you have a wff that begins with "~" and then a modal operator; this moves an operator to the beginning of the formula, so we can drop it later.

We drop modal operators using the next two rules (which hold regardless of what wff replaces "A"). Here's the drop-diamond rule:

Drop diamond DD
◇A → W ∴ A, use a *new* string of W's

Here the line with "◇A" can use any world prefix – and the line with "∴ A" must use a *new* string of one or more W's (a string not occurring in earlier lines). If "A" is possible, then "A" is true in *some* possible world; we can give this world a name – but a *new* name, since "A" needn't be true in any of the worlds used in the proof so far. We'll use "W" for the first diamond we drop, "WW" for the second, and so forth. So if we drop two diamonds, then we introduce two new worlds:

◇H
◇T______
W ∴ H
WW ∴ T

Heads is possible, tails is possible; let's call an imagined world with heads "W," and one with tails "WW." It's OK to use "W" in the first inference, since it occurs in no earlier line. But the second inference must use "WW," since "W" has now already occurred.

We can drop diamonds from longer formulas, if the diamond *begins* the wff. So this first inference is fine:

◇(A • B)__
W ∴ (A • B)

These next two examples are wrong (since the formula doesn't begin with a diamond – instead, it begins with a left-hand parenthesis):

(◇A ⊃ B)
―――――――――
W ∴ (A ⊃ B)

(◇A • ◇B)
―――――――――
W ∴ (A • B)

Drop only *initial operators* (diamonds or boxes).

Here's the drop-box rule: 0236

Drop box DB
☐A → W ∴ A, use any world prefix

The lines with "☐A" and "∴ A" can use *any* world prefixes, the same or different, including the blank world prefix for the actual world. If "A" is necessary, then "A" is true in *all* possible worlds, and so we can put "A" in any world. But it's bad strategy to drop a box into a *new* world; stay in *old* worlds. As before, we can drop boxes from longer formulas, as long as the box *begins* the wff. So this next inference is fine:

☐(A ⊃ B)
―――――――――
W ∴ (A ⊃ B)

These next two example are wrong (since the formula doesn't *begin* with a box – instead it begins with a left-hand parenthesis – *drop only initial operators*):

(☐A ⊃ B)
―――――――――
W ∴ (A ⊃ B)

(☐A ⊃ ☐B)
―――――――――
W ∴ (A ⊃ B)

"(☐A ⊃ B)" and "(☐A ⊃ ☐B)" are if-then forms and follow the if-then rules: if we have the first part true, we can get the second true; if we have the second part false, we can get the first false; if we get stuck, we make an assumption.

Here's a valid modal argument and its proof:

Necessarily, if there's rain then there's precipitation.
It's possible that there's rain.
∴ It's possible that there's precipitation.

```
  1  ☐(R ⊃ P)    Valid
* 2  ◇R
  [ ∴ ◇P
* 3 ┌ asm: ~◇P
  4 │ ∴ ☐~P    {from 3}
  5 │ W ∴ R    {from 2}
* 6 │ W ∴ (R ⊃ P)    {from 1}
  7 │ W ∴ P    {from 5 and 6}
  8 └ W ∴ ~P    {from 4}
  9 ∴ ◇P    {from 3; 7 contradicts 8}
```

Assume "It's not possible that there's precipitation." Reverse the squiggle to get "It's necessary that there's no precipitation" in 4. Drop the diamond in 2, using a new world W, to get "There's rain" in W in 5. Drop the box in 1 to get "If there's rain then there's precipitation" in W in 6. From these two, get "There's precipitation" in W in 7. Drop a box again in 8 to get contradiction. Our conclusion follows: "It's possible that there's precipitation."

We'll typically use modal rules in this order: (1) First reverse squiggles; (2) then drop initial diamonds, using a new world each time; (3) lastly, drop each initial box once for each old world. Star when reversing squiggles or dropping a diamond (starred lines have redundant information and can largely can be ignored in deriving further lines):

```
*  ~☐A
─────────
 ∴ ◇~A
```

```
*  ◇A
────────
 W ∴ A
```

0237 *Don't star when dropping a box*; we can never exhaust a "necessary" statement – and we may have to use it again later in the proof.

Here's an easy modal proof:

```
  1  ☐(A • B)    Valid
  [ ∴ ☐A
* 2 ┌ asm: ~☐A
* 3 │ ∴ ◇~A    {from 2}
  4 │ W ∴ ~A    {from 3}
```

5 │ W ∴ (A • B) {from 1}
6 └ **W ∴ A {from 5}**
7 ∴ ☐A {from 2; 4 contradicts 6}

Reverse squiggles to get "◇~A" in line 3. Drop a diamond to get "W ∴ ~A" in line 4. Then drop a box to get "W ∴ (A • B)" in line 5.

In this proof, there's no point to dropping the box into the actual world, to go from "☐(A • B)" in line 1 to "∴ (A • B)" with no initial W's. Drop a box into the actual world (besides into any W-worlds) only in these two cases:

1. The original premises or conclusion have an *unmodalized* instance of a letter. (A letter is unmodalized if it doesn't occur as part a larger wff beginning with "☐" or "◇"; in "(A • ◇A)" only the first "A" is *unmodalized*.)

2. You've done everything else possible (including further assumptions if needed) and still have no other old worlds.

Here are examples:

Case 1: unmodalized letter

1 ☐(A • B)
[∴ A ⇐ *unmodalized*
2 ┌ **asm: ~A**
3 │ ∴ (A • B) {from 1}
4 └ **∴ A {from 3}**
5 ∴ A {from 2; 2 contradicts 4}

Here the original argument has an unmodalized letter. When this happens, drop boxes into the actual world (as in line 3) and also into all W-worlds (if there are any).

Case 2: no other worlds

1 ☐~A
[∴ ~☐A
2 ┌ asm: ☐A
3 │ **∴ A {from 2}**
4 └ **∴ ~A {from 1}**
5 ∴ ~☐A {from 2; 3 contradicts 4}

Here, when you drop the box to get line 3, there are no other old worlds (since you had no diamonds to drop); so use the actual world (with no W's).

Our "standard strategy" here has us drop boxes into the actual world in these two cases, and only these. This always works, but it sometimes gives us lines that we don't need; we can skip these lines if we see that we don't need them.
0238
In doing proofs, first assume the conclusion's opposite; then use modal rules plus S- and I-rules to derive all you can. If you find a contradiction, apply RAA. If you're stuck and need to break a NOT-BOTH, OR, or IF-THEN, then make another assumption. If you get no contradiction and yet can't do anything further, then try to refute. Here's a fuller statement of our strategy's modal steps:

1. FIRST REVERSE SQUIGGLES: For each unstarred, not-blocked-off line that begins with "∼☐" or "∼◇," derive a line using the reverse-squiggle rules. Star the original line.

2. THEN DROP DIAMONDS: For each unstarred, not-blocked-off line that begins with a diamond, derive an instance using the next available new world (but don't drop a diamond if you already have a not-blocked-off-instance in some previous line – so don't drop "◇A" if you already have "W ∴ A"). Star the original line.

3. LASTLY DROP BOXES: For each not-blocked-off line that begins with a box, derive instances using each old world. Don't star the original line; you might have to use it again. (Drop boxes into the actual world under the two conditions given on the previous page.)

Drop diamonds before boxes. Introduce a new world each time you drop a diamond, and use the same old worlds when you drop a box. And drop only initial diamonds and boxes.

10.2a Exercise: LogiCola KV

Prove each of these arguments to be valid (all are valid).

☐(A ⊃ B) ◇∼B ∴ ◇∼A

1 □(A ⊃ B) **Valid**
* 2 ◇~B
[∴ ◇~A
* 3 asm: ~◇~A
4 ∴ □A {from 3}
5 **W ∴ ~B {from 2}**
* 6 W ∴ (A ⊃ B) {from 1}
7 W ∴ A {from 4}
8 **W ∴ B {from 6 and 7}**
9 ∴ ◇~A {from 3; 5 contradicts 8}

1. ◇(A • B)
∴ ◇A

2. A
∴ ◇A

3. ~◇(A • ~B)
∴ □(A ⊃ B)

4. □(A ∨ ~B)
~□A
∴ ◇~B

5. (◇A ∨ ◇B)
∴ ◇(A ∨ B)

6. (A ⊃ □B)
◇~B
∴ ◇~A

7. ~◇(A • B)
◇A
∴ ~□B 0239

8. □A
∴ ◇A

9. ☐A
~☐B
∴ ~☐(A ⊃ B)

10. ☐(A ⊃ B)
∴ (☐A ⊃ ☐B)

10.2b Exercises: LogiCola KV

First appraise intuitively. Then translate into logic (using the letters given) and prove to be valid (all are valid).

1. "You knowingly testify falsely because of threats to your life" entails "You lie."
It's possible that you knowingly testify falsely because of threats to your life but don't intend to deceive. (Maybe you hope no one will believe you.)
∴ "You lie" is consistent with "You don't intend to deceive." [Use T, L, and I; from Tom Carson, who writes on the morality of lying.]

2. Necessarily, if you don't decide then you decide not to decide.
Necessarily, if you decide not to decide then you decide.
∴ Necessarily, if you don't decide then you decide. [Use D for "You decide" and N for "You decide not to decide." This is adapted from Jean-Paul Sartre.]

3. If truth is a correspondence with the mind, then "There are truths" entails "There are minds."
"There are minds" isn't logically necessary.
Necessarily, if there are no truths then it is not true that there are no truths.
∴ Truth isn't a correspondence with the mind. [Use C, T, and M.]

4. There's a perfect God.
There's evil in the world.
∴ "There's a perfect God" is logically compatible with "There's evil in the world." [Use G and E. Most who doubt the conclusion would also doubt premise 1.]

5. "There's a perfect God" is logically compatible with T.
T logically entails "There's evil in the world."
∴ "There's a perfect God" is logically compatible with "There's evil in the world." [Use G, T, and E. Here T (for "theodicy") is a possible explanation of why God permits evil that's consistent with God's perfection and entails the existence of evil. T might say: "The world has evil because God, who is perfect, wants us to make significant free choices to struggle to bring a half-completed world toward its fulfillment; moral evil comes from the abuse of human freedom and physical evil from the half-completed state of the world." This basic argument (but not the specific T) is from Alvin Plantinga.]

6. “There’s a perfect God and there’s evil in the world and God has some reason for permitting the evil” is logically consistent.
∴ “There’s a perfect God and there’s evil in the world” is logically consistent. [Use G, E, and R. This is Ravi Zacharias’s version of Plantinga’s argument.] 0240

7. God is omnipotent.
“You freely always do the right thing” is logically possible.
If “You freely always do the right thing” is logically possible and God is omnipotent, then it’s possible for God to bring it about that you freely always do the right thing.
∴ It’s possible for God to bring it about that you freely always do the right thing. [Use O, F, and B; from J. L. Mackie. He thought God had a third option besides making robots who always act rightly and free beings who sometimes act wrongly: he could make free beings who always act rightly.]

8. “God brings it about that you do A” is inconsistent with “You freely do A.”
“God brings it about that you freely do A” entails “God brings it about that you do A.”
“God brings it about that you freely do A” entails “You freely do A.”
∴ It’s impossible for God to bring it about that you freely do A. [Use B, F, and G. This attacks the conclusion of the previous argument.]

9. “This is a square” entails “This is composed of straight lines.”
“This is a circle” entails “This isn’t composed of straight lines.”
∴ “This is a square and also a circle” is self-contradictory. [S, L, C]

10. “This is red and there’s a blue light that makes red things look violet to normal observers” entails “Normal observers won’t sense redness.”
“This is red and there’s a blue light that makes red things look violet to normal observers” is logically consistent.
∴ “This is red” doesn’t entail “Normal observers will sense redness.” [Use R, B, and N; from Roderick Chisholm.]

11. “All brown dogs are brown” is a necessary truth.
“Some dog is brown” isn’t a necessary truth.
“Some brown dog is brown” entails “Some dog is brown.”
∴ “All brown dogs are brown” doesn’t entail “Some brown dog is brown.” [Use A for “All brown dogs are brown,” X for “Some dog is brown,” and S for “Some brown dog is brown.” This attacks a doctrine of traditional logic (§2.8), that “all A is B” entails “some A is B.”]

12. It's necessary that, if God exists as a possibility but does not exist in reality, then there could be a being greater than God (namely, a similar being that also exists in reality).
"There could be a being greater than God" is self-contradictory (since "God" is defined as "a being than which no greater could be").
It's necessary that God exists as a possibility.
∴ It's necessary that God exists in reality. [Use P for "God exists as a possibility," R for "God exists in reality," and G for "There's a being greater than God." This is a modal version of St Anselm's ontological argument.] 0241

13. If "X is good" and "I like X" are interchangeable, then "I like hurting people" logically entails "Hurting people is good."
"I like hurting people but hurting people isn't good" is consistent.
∴ "X is good" and "I like X" aren't interchangeable. [Use I, L, and G. This argument attacks subjectivism.]

14. "You sin" entails "You know what you ought to do and you're able to do it and you don't do it."
It's necessary that if you know what you ought to do then you want to do it.
It's necessary that if you want to do it and you're able to do it then you do it.
∴ It's impossible for you to sin. [S, K, A, D, W]

15. Necessarily, if it's true that there are no truths then there are truths.
∴ It's necessary that there are truths. [Use T for "There are truths."]

10.3 Refutations

Applying our proof strategy to an invalid argument leads to a refutation:

It's possible that it's heads.
It's possible that it's tails.
∴ It's possible that it's both heads and tails.

```
* 1  ◇H    Invalid
* 2  ◇T
     [ ∴ ◇(H • T)
* 3  asm: ~◇(H • T)
  4  ∴ □~(H • T)    {from 3}
  5  W ∴ H    {from 1}
  6  WW ∴ T    {from 2}
* 7  W ∴ ~(H • T)    {from 4}
* 8  WW ∴ ~(H • T)    {from 4}
  9  W ∴ ~T    {from 5 and 7}
```

10 WW ∴ ~H {from 6 and 8}

W	H, ~T
WW	T, ~H

Reverse a squiggle (line 4). Drop two diamonds, using a new world each time (lines 5 and 6). Drop the box twice, using W and WW (lines 7 and 8). Getting no contradiction, we gather simple wffs for a *refutation*. We get a little galaxy of two possible worlds: one with heads-and-not-tails and another with tails-and-not-heads. The argument is invalid, since this galaxy makes the premises both true (since it's heads in one possible world and tails in another) but the conclusion false (since no possible world has both heads and tails).

If we try to prove an invalid argument, we'll instead be led to a refutation – a galaxy of possible worlds that make the premises all true and conclusion false. In evaluating premises and conclusion, use these rules to evaluate each formula or subformula that starts with a modal operator: 0242

"◇A" is true if and only if *at least one world* has "A" true.
"☐A" is true if and only if *all worlds* have "A" true.

Premise "◇H" is true because world W has "H" true, and premise "◇T" is true because world WW has "T" true.[1] But conclusion "◇(H • T)" is false because no world has "(H • T)" true:

In W: (H • T) = (1 • 0) = 0

In WW: (H • T) = (0 • 1) = 0

Always check that your refutation works. If you don't get premises all 1 and conclusion 0, then you did something wrong; look at what you did with the wff that came out wrong (a premise that's 0 or ?, or a conclusion that's 1 or ?).

These two rules are crucial for working out proofs and refutations:

- For each initial *diamond*, introduce a *new world*.
- For each initial *box*, derive an instance for each *old world*.

[1] POSSIBLE is like OR: something holds in this world OR that world OR that world … – so a single true case makes a POSSIBLE true. NECESSARY is like AND: something holds in this world AND that world AND that world … – so a single false case makes a NECESSARY false.

If you have two diamonds, don't drop both using the same world – and don't drop just one diamond. And if you have two worlds, then drop any box using *both* worlds; if in our example we dropped the box in "☐∼(H • T)" using "W" but not "WW," then our attempted refutation would fail:

W	H, ∼T
WW	T

W has heads and not tails, WW has tails

Since "H" is unknown in WW, our conclusion "◇(H • T)" would also be unknown (because the second case with "WW" is unknown):

$$\text{In W: } (H \bullet T) = (1 \bullet 0) = 0$$

$$\text{In WW: } (\mathbf{H} \bullet T) = (\mathbf{?} \bullet 1) = \mathbf{?}$$

The "It's possible that it's both heads and tails" conclusion is unknown, since our world WW doesn't exclude it being heads (besides being tails). We avoid such problems if we drop each initial box using each old world; here we'd go from "☐∼(H • T)" to "WW ∴ ∼(H • T)," which would lead to "WW ∴ ∼H."

As we refute arguments, we'll often have to evaluate premises or conclusions that don't *start* with boxes or diamonds, such as these wffs: 0243

∼☐H
∼☐H

∼☐(H ∨ T)
∼☐(H ∨ T)

∼◇(H • T)
∼◇(H • T)

Identify any *subformulas* that start with a boxes or diamonds (as highlighted here). Evaluate each subformula to be 1 or 0, and then apply "∼" to reverse the result. On our heads-tails refutation, "☐H" = 0, and so "∼☐H" = 1. Likewise, "☐(H ∨ T)" = 1, and so "∼☐(H ∨ T)" = 0; and "◇(H • T)" = 0, and so "∼◇(H • T)" = 1. In evaluating a wff that starts with a squiggle and then a box-or-diamond, evaluate the wff without the squiggle and then give the original wff the opposite value. Divide and conquer!

Here's another invalid argument:

1 (☐A ⊃ ☐B) **Invalid**
[∴ (A ⊃ B)
* 2 asm: ∼(A ⊃ B)
3 ∴ A {from 2}
4 ∴ ∼B {from 2}
** 5 asm: ∼☐A {break 1}
** 6 ∴ ◇∼A {from 5}
7 W ∴ ∼A {from 6}

	A, ∼B
W	∼A

Our refutation has an actual world and a possible world W. To evaluate the premise, first identity and evaluate subformulas that start with a box or diamond (these are highlighted here), and then plug in 1 or 0 for these:

For "(☐A ⊃ ☐B)," we first evaluate "☐A" and "☐B."
"☐A" is false because "A" is false in W.
"☐B" is false because "B" is false in the actual world.
Replace both with "0."
We get "(0 ⊃ 0)," which simplifies to "1."
So "(☐A ⊃ ☐B)" is true.

The conclusion is "(A ⊃ B)," which uses unmodalized letters; these should be evaluated in the actual world. So conclusion (A ⊃ B) = (1 ⊃ 0) = 0. Since we have true premises and a false conclusion, our argument is invalid.

As we refute invalid arguments, we'll often have complex premises or conclusions to evaluate, such as these wffs:

(☐A ⊃ ☐B)
(☐A ⊃ ☐B)

(☐(F ∨ G) ⊃ (∼◇G • ◇∼H))
(☐(F ∨ G) ⊃ (∼◇G • ◇∼H))

(∼☐(F ⊃ G) ≡ ∼◇H)
(∼☐(F ⊃ G) ≡ ∼◇H)

As above, first identity *subformulas* that start with boxes or diamonds (as highlighted). Evaluate each such subformula to be 1 or 0, replace it with 1 or 0, and figure out whether the whole formula is 1 or 0. Divide and conquer!

0244

This English argument has an ambiguous first premise, which could have two different meanings:

If you're a bachelor, then you must be unmarried.
You're a bachelor.
∴ It's logically necessary that you're unmarried.

(B ⊃ ☐U) If you're a bachelor, then you're *inherently unmarriable*.

☐(B ⊃ U) It's necessary that *if* you're a bachelor *then* you're unmarried.

Work out both versions:

Box-inside version (valid but premise 1 is false):

* 1 (B ⊃ ☐U) **Valid**
2 B
[∴ ☐U
3 ┌ asm: ~☐U
4 └ ∴ ☐U {from 1 and 2}
5 ∴ ☐U {from 3; 3 contradicts 4}

Box-outside version (invalid):

1 ☐(B ⊃ U) **Invalid**
2 B
[∴ ☐U
* 3 asm: ~☐U
* 4 ∴ ◇~U {from 3}
5 W ∴ ~U {from 4}
* 6 W ∴ (B ⊃ U) {from 1}
* 7 ∴ (B ⊃ U) {from 1}
8 W ∴ ~B {from 5 and 6}
9 ∴ U {from 2 and 7}

	B, U
W	~B, ~U

Both versions are flawed: the first has a false premise, while the second is invalid. So the proof that you're inherently unmarriable fails. Arguments with a modal ambiguity often have one interpretation with a false premise and another that's invalid; such arguments often seem sound until we focus on the ambiguity.

10.3a Exercise: LogiCola KI

Prove each of these arguments to be invalid (all are invalid).

☐(A ⊃ B) ◇A ∴ ☐B

1 ☐(A ⊃ B) **Invalid**
* 2 ◇A
   [∴ ☐B
* 3 asm: ~☐B
* 4 ∴ ◇~B {from 3}
   5 W ∴ ~B {from 4}
   6 WW ∴ A {from 2}
* 7 W ∴ (A ⊃ B) {from 1}
* 8 WW ∴ (A ⊃ B) {from 1}
   9 W ∴ ~A {from 5 and 7}
  10 WW ∴ B {from 6 and 8}

W	~A, ~B
WW	A, B

0245

1. ◇A
∴ ☐A

2. A
∴ ☐A

3. ◇A
◇B
∴ ◇(A • B)

4. ☐(A ⊃ ~B)
B
∴ ☐~A

5. $(\Box A \supset \Box B)$
$\therefore \Box(A \supset B)$

6. $\Diamond A$
$\sim\Box B$
$\therefore \sim\Box(A \supset B)$

7. $\Box(C \supset (A \vee B))$
$(\sim A \bullet \Diamond\sim B)$
$\therefore \Diamond\sim C$

8. $\Box(A \vee \sim B)$
$\therefore (\sim\Diamond B \vee \Box A)$

9. $\Box((A \bullet B) \supset C)$
$\Diamond A$
$\Diamond B$
$\therefore \Diamond C$

10. $\sim\Box A$
$\Box(B \equiv A)$
$\therefore \sim\Diamond B$

10.3b Exercise: LogiCola KC

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation). Translate ambiguous English arguments both ways; prove or disprove each symbolization separately.

1. If the pragmatist view of truth is right, then “A is true” entails “A is useful to believe.”
“A is true but not useful to believe” is consistent.
∴ The pragmatist view of truth isn’t right. [Use P, T, and B.]

2. You know.
“You’re mistaken” is logically possible.
∴ “You know and are mistaken” is logically possible. [Use K and M.]

3. Necessarily, if this will be then this will be.
∴ If this will be, then it’s necessary (in itself) that this will be. [Use B. This illustrates two senses of “Que será será” – “Whatever will be will be.” The first sense is a truth of logic while the second is a form of fatalism.]

4. I’m still.
If I’m still, then it’s necessary that I’m not moving.
If it’s necessary that I’m not moving, then whether I move is not a matter of my free choice.
∴ Whether I move is not a matter of my free choice. [Use S, M, and F. This is adapted from the medieval thinker Boethius, who used a similar example to explain the box-inside/box-outside distinction.]

5. It’s necessarily true that if you’re morally responsible for your actions then you’re free.
It’s necessarily true that if your actions are uncaused then you aren’t morally responsible for your actions.
∴ “You’re free” doesn’t entail “Your actions are uncaused.” [Use R, F, and U; from A. J. Ayer.] 0246

6. If “One’s conscious life won’t continue forever” entails “Life is meaningless,” then a finite span of life is meaningless.
If a finite span of life is meaningless, then an infinite span of life is meaningless.
If an infinite span of life is meaningless, then “One’s conscious life will continue forever” entails “Life is meaningless.”
∴ If it’s possible that life is not meaningless, then “One’s conscious life won’t continue forever” doesn’t entail “Life is meaningless.” [C, L, F, I]

7. If you have money, then you couldn’t be broke.
You could be broke.
∴ You don’t have money. [Use M and B. Is this argument just a valid instance of *modus tollens*: “$(P \supset Q)$, $\sim Q \therefore \sim P$”?]

8. If you know, then you couldn’t be mistaken.
You could be mistaken.
∴ You don’t know. [Use K and M. Since we could repeat this reasoning for any alleged item of knowledge, the argument seems to show that genuine knowledge is impossible.]

9. It’s necessary that if there’s a necessary being then “There’s a necessary being” (by itself) is necessary.
“There’s a necessary being” is logically possible.
∴ “There’s a necessary being” is logically necessary. [Use N for “There’s a necessary being” or “There’s a being that exists of logical necessity”; this being is often identified with God; from Charles Hartshorne and St Anselm; it’s sometimes called “Anselm’s second ontological argument.” The proof raises logical issues that we’ll deal with in the next chapter.]

10. It's necessary that either I'll do it or I won't do it.
If it's necessary that I'll do it, then I'm not free.
If it's necessary that I won't do it, then I'm not free.
∴ I'm not free.   [Use D for "I'll do it." Aristotle and the Stoic Chrysippus discussed this argument. This argument's flaw relates to a point made by Chrysippus, that "$\Box(D \vee \sim D) \therefore (\Box D \vee \Box \sim D)$" is invalid and is like arguing "Everything is either A or non-A; therefore either everything is A or everything is non-A."]

11. "This agent's actions were all determined" is consistent with "I describe this agent's character in an approving way."
"I describe this agent's character in an approving way" is consistent with "I praise this agent."
∴ "This agent's actions were all determined" is consistent with "I praise this agent."   [D, A, P]

12. If thinking is just a chemical brain process, then "I think" entails "There's a chemical process in my brain."
"There's a chemical process in my brain" entails "I have a body."
"I think but I don't have a body" is logically consistent.
∴ Thinking isn't just a chemical brain process.   [Use J, T, C, and B. This argument attacks a form of materialism.] 0247

13. If "I did that on purpose" entails "I made a prior purposeful decision to do that," then there's an infinite chain of previous decisions to decide.
It's impossible for there to be an infinite chain of previous decisions to decide.
∴ "I did that on purpose" is consistent with "I didn't make a prior purposeful decision to do that."   [Use D, P, and I; from Gilbert Ryle.]

14. God knew that you'd do it.
If God knew that you'd do it, then it was necessary that you'd do it.
If it was necessary that you'd do it, then you weren't free.
∴ You weren't free.   [Use K, D, and F. This argument is the focus of an ancient controversy. Would divine foreknowledge preclude human freedom? If it would, then should we reject human freedom (as did Luther) or divine foreknowledge (as did Charles Hartshorne)? Or perhaps (as the medieval thinkers Boethius, Aquinas, and Ockham claimed) is there a flaw in the argument that divine foreknowledge precludes human freedom?]

15. If "good" means "socially approved," then "Racism is socially approved" logically entails "Racism is good."
"Racism is socially approved but not good" is consistent.
∴ "Good" doesn't mean "socially approved." [Use M, S, and G. This argument attacks cultural relativism.]

16. Necessarily, if God brings it about that A is true, then A is true.
A is a self-contradiction.
∴ It's impossible for God to bring it about that A is true.   [Use B and A, where B is for "God brings it about that A is true."]

17. If this is experienced, then this must be thought about.
"This is thought about" entails "This is put into the categories of judgments."
∴ If it's possible for this to be experienced, then it's possible for this to be put into the categories of judgments.   [Use E, T, and C; from Immanuel Kant, who argued that our mental categories apply, not necessarily to everything that exists, but rather to everything that we could experience.]

18. Necessarily, if formula B has an all-1 truth table then B is true.
∴ If formula B has an all-1 truth table, then B (taken by itself) is necessary.   [Use A and B. This illustrates the box-outside versus box-inside distinction.]

19. Necessarily, if you mistakenly think that you exist then you don't exist.
Necessarily, if you mistakenly think that you exist then you exist.
∴ "You mistakenly think that you exist" is impossible.   [Use M and E. This relates to Descartes's "I think, therefore I am" ("*Cogito ergo sum*").]

20. If "good" means "desired by God," then "This is good" entails "There's a God."
"There's no God, but this is good" is consistent.
∴ "Good" doesn't mean "desired by God."   [Use M, A, and B. This attacks one form of the divine command theory of ethics. Some (see 9 and 26 of this section and 12 of §10.2b) say, against premise 2, that "There's no God" is logically impossible.]
0248

21. If Plato is right, then it's necessary that ideas are superior to material things.
It's possible that ideas aren't superior to material things.
∴ Plato isn't right.   [P, S]

22. "I seem to see a chair" doesn't entail "There's an actual chair that I seem to see."
If we directly perceive material objects, then "I seem to see a chair and there's an actual chair that I seem to see" is consistent.
∴ We don't directly perceive material objects.   [S, A, D]

23. "There's a God" is logically incompatible with "There's evil in the world."
There's evil in the world.
∴ "There's a God" is self-contradictory.   [G, E]

24. If you do all your homework right, then it's impossible that you get this problem wrong.
It's possible that you get this problem wrong.
∴ You don't do all your homework right. [R, W]

25. "You do what you want" is compatible with "Your act is determined."
"You do what you want" entails "Your act is free."
∴ "Your act is free" is compatible with "Your act is determined." [W, D, F]

26. It's necessarily true that if God doesn't exist in reality then there's a being greater than God (since then any existing being would be greater than God).
It's not possible that there's a being greater than God (since "God" is defined as "a being than which no being could be greater").
∴ It's necessary that God exists in reality. [Use R and B. This is a simplified modal form of St Anselm's ontological argument.]

27. It was always true that you'd do it.
If it was always true that you'd do it, then it was necessary that you'd do it.
If it was necessary that you'd do it, then you weren't free.
∴ You weren't free. [Use A (for "It was always true that you'd do it" – don't use a box here), D, and F. This argument is much like problem 14. Are statements about future contingencies (for example, "I'll brush my teeth tomorrow") true or false before they happen? Should we do truth tables for such statements in the normal way, assigning them "1" or "0"? Does this preclude human freedom? If so, should we then reject human freedom? Or should we adopt a many-valued logic that says that statements about future contingencies aren't "1" or "0" but must instead have some third truth value (maybe "½")? Or is the argument fallacious?]

11 Further Modal Systems

Modal logic studies arguments whose validity depends on "possible," "necessary," and similar notions. The previous chapter presented a basic system that builds on propositional logic. This present chapter considers alternative systems of propositional and quantified modal logic.

11.1 Galactic travel

While logicians usually agree on which arguments are valid, there are more disagreements about modal arguments. Many disputes involve arguments in which one modal operator occurs within the scope of another – like "◇◇A ∴ ◇A" and "☐(A ⊃ ☐B), ◇A ∴ B." These disputes reflect differences in how to formulate the drop-box rule. So far, we've assumed a system called "S5," which lets us go from any world to any world when we drop a box (§10.2):

Drop box DB ☐A → W ∴ A, use any world prefix
Here the line with "☐A" and the line with "W ∴ A" can use any world prefixes, the same or different.

This assumes that whatever is necessary in *any* world is thereby true in *all* worlds without restriction. A further implication is that whatever is necessary in one world is thereby necessary in all worlds.

Some weaker views reject these ideas. On these views, what is necessary only has to be true in all "suitably related" worlds; so these views restrict the drop-box rule. All the views in question let us go from "☐A" in a world to "A" in the *same* world. But we can't always go from "☐A" in one world to "A" in *another* world; traveling between worlds requires (at least on my way of expressing it) a suitable "travel ticket."

We get travel tickets when we drop diamonds. Let "W1" and "W2" stand for world prefixes. Suppose we go from "◇A" in world W1 to "A" in new world W2. Then we get a travel ticket from W1 to W2, and we'll write "W1 ⇒

W2”:

W1 ⇒ W2
We have a ticket to move from world W1 to world W2

0250 Suppose we do a proof with wffs “◇◇A” and “◇B.” We’d get these travel tickets when we drop diamonds (here “#” stands for the actual world):

1 ◇◇A
2 ◇B
....
11 W ∴ ◇A {from 1} **# ⇒ W**
12 WW ∴ A {from 11} **W ⇒ WW**
13 WWW ∴ B {from 2} **# ⇒ WWW**

Dropping a diamond gives us a travel ticket from the world in the “from” line to the world in the “to” line. So in line 11 we get ticket “# ⇒ W” – because we moved from “◇◇A” in the actual world (“#”) to “◇A” in world W. Tickets are reusable; we can use “W1 ⇒ W2” any number of times.

The rules for using tickets vary. System T lets us use only one ticket at a time, and only in the direction of the arrow; system S4 lets us combine a series of tickets, while system B lets us use them in a backwards direction. Suppose we have “☐A” in world W1 and want to put “A” in world W2:

- *System T.* We need a *ticket* from W1 to W2.
- *System S4*. Like T, but we also can use a *series* of tickets.
- *System B*. Like T, but a ticket also works *backwards*.

Suppose we have three travel tickets:

⇒ W
W ⇒ WW
⇒ WWW

System T would let us, when we drop boxes, go from # to W, from W to WW, and from # to WWW. The other systems allow these and more. System S4 lets us use a *series* of tickets in the direction of the arrow; this lets us go from # to WW. System B lets us use single tickets *backwards*; this lets us go from W to #, from WW to W, and from WWW to #. In contrast, system S5 lets us go from any world to any world; this is equivalent to letting us use any ticket or series of tickets in either direction.

S5 is the most liberal system and accepts the most valid arguments; so S5 is the strongest system. T is the weakest system, allowing the fewest proofs. S4 and B are intermediate, each allowing some proofs that the other doesn't. The four systems give the same result for most arguments. But some arguments are valid in one system but invalid in another; these arguments use wffs that apply a modal operator to a wff already containing a modal operator.

This argument is valid in S4 or S5 but invalid in T or B: 0251

```
    1   ☐A
    [ ∴ ☐☐A
  * 2 ┌ asm: ~☐☐A
  * 3 │ ∴ ◇~☐A    {from 2}
  * 4 │ W ∴ ~☐A    {from 3}    # ⇒ W
  * 5 │ W ∴ ◇~A    {from 4}
    6 │ WW ∴ ~A    {from 5}    W ⇒ WW
    7 └ WW ∴ A    {from 1}    Need S4 or S5
    8 ∴ ☐☐A    {from 2; 6 contradicts 7}
```

Line 7 requires that we combine a series of tickets in the direction of the arrow. Tickets "# ⇒ W" and "W ⇒ WW" then let us go from actual world # (line 1) to world WW (line 7). This requires systems S4 or S5.

This next one is valid in B or S5 but invalid in T or S4:

```
    1   A
    [ ∴ ☐◇A
  * 2 ┌ asm: ~☐◇A
  * 3 │ ∴ ◇~◇A    {from 2}
  * 4 │ W ∴ ~◇A    {from 3}    # ⇒ W
    5 │ W ∴ ☐~A    {from 4}
    6 └ ∴ ~A    {from 5}    Need B or S5
    7 ∴ ☐◇A    {from 2; 1 contradicts 6}
```

Line 6 requires using ticket "# ⇒ W" backwards, to go from world W (line 5) to the actual world # (line 6). This requires systems B or S5.

This last one is valid in S5 but invalid in T or B or S4:

```
  * 1   ◇A
    [ ∴ ☐◇A
  * 2 ┌ asm: ~☐◇A
```

* 3 | ∴ ◇∼◇A {from 2}
* 4 | W ∴ ∼◇A {from 3} **# ⇒ W**
5 | W ∴ ☐∼A {from 4}
6 | **WW ∴ A {from 1} # ⇒ WW**
7 └ **WW ∴ ∼A {from 5} Need S5**
8 ∴ ☐◇A {from 2; 6 contradicts 7}

Line 7 requires combining a series of tickets and using some backwards. Tickets "# ⇒ W" and "# ⇒ WW" then let us go from W (line 5) to WW (line 7). This requires system S5.

S5 is the simplest system in several ways:

- We can formulate S5 more simply. The box-dropping rule doesn't have to mention travel tickets; we need only say that, if we have "☐A" in any world, then we can put "A" in any world (the same or a different one). 0252
- S5 captures simple intuitions about necessity and possibility: what's necessary is what's true in all worlds, what's possible is what's true in some worlds, and what's necessary or possible doesn't vary between worlds.
- On S5, any string of boxes and diamonds simplifies to its last symbol. So "☐☐" and "◇☐" simplify to "☐," and "◇◇" and "☐◇" simplify to "◇."

Which is the best system? This depends on what we take the box and diamond to mean. If we take them to be about the logical necessity and possibility of *ideas*, then S5 is the best system. If an idea (for example, the claim that 2 = 2) is logically necessary, then it couldn't have been other than logically necessary. So if A is logically necessary, then it's logically necessary that A is logically necessary ["(☐A ⊃ ☐☐A)"]. Similarly, if an idea is logically possible, then it's logically necessary that it's logically possible ["(◇A ⊃ ☐◇A)"]. Of the four systems, only S5 accepts both formulas. All this presupposes that we use the box to talk about the logical necessity of *ideas*.

Or we could take the box to be about the logical necessity of *sentences*. The *sentence* "2 = 2" just happens to express a necessary truth; it wouldn't have expressed one if English had used "=" to mean "≠." So the *sentence* is necessary, but it's not necessary that it's necessary; this makes "(☐A ⊃ ☐☐A)" false. The *idea* that "2 = 2" now expresses, however, is both necessary and necessarily necessary; a change in our language wouldn't make this *idea* false, but it would change how we'd express this idea. So whether S5 is the best system can depend on whether we take the box to be about the necessity of *ideas* or of *sentences*.

There are still other ways to take "necessary." Sometimes calling something "necessary" might mean that it's "physically necessary," "proven," "known," or "obligatory." Some logicians like the weak system T because it holds for various senses of "necessary"; such logicians might still use S5 for arguments about the logical necessity of *ideas*. While I have sympathy with this view, most of the modal arguments I'm interested in are about the logical necessity of ideas. So I use S5 as the standard system of modal logic but feel free to switch to weaker systems for arguments about other kinds of necessity.

Here we've considered the four main modal systems. We could invent other systems – for example, ones in which we can combine travel tickets only in groups of three. Logicians develop such systems, not to help us in analyzing real arguments, but rather to explore interesting formal structures.[1]

11.1a Exercise: LogiCola KG

Using system S5, prove each of these arguments to be valid. Also say in which systems the argument is valid: T, B, S4, or S5. 0253

~☐A ∴ ☐~☐A
* 1 ~☐A **Valid** [∴ ☐~☐A * 2 ┌ asm: ~☐~☐A * 3 │ ∴ ◇☐A {from 2} * 4 │ ∴ ◇~A {from 1} 5 │ W ∴ ☐A {from 3} **# ⇒ W** 6 │ **WW ∴ ~A {from 4} # ⇒ WW** 7 └ **WW ∴ A {from 5} Need S5** 8 ∴ ☐~☐A {from 2; 6 contradicts 7} Line 7 combines a series of tickets and uses some backwards. This requires S5.

1. ◇☐A
∴ A

2. ◇A
∴ ◇◇A

3. ◇◇A
∴ ◇A

[1] For more on alternative modal systems, consult G. E. Hughes and M. J. Cresswell, *A New Introduction to Modal Logic* (London: Routledge, 1996).

4. $\Diamond\Box A$
$\therefore \Box A$

5. $(\Box A \supset \Box B)$
$\therefore \Box(\Box A \supset \Box B)$

6. $\Box(A \supset B)$
$\therefore \Box(\Box A \supset \Box B)$

7. $(\Diamond A \supset \Box B)$
$\therefore \Box(A \supset \Box B)$

8. $\Box(A \supset \Box B)$
$\therefore (\Diamond A \supset \Box B)$

9. $\Diamond\Box\Diamond A$
$\therefore \Diamond A$

10. $\Diamond A$
$\therefore \Diamond\Box\Diamond A$

11. $\Box A$
$\therefore \Box(B \supset \Box A)$

12. $\Box\Diamond\Box\Diamond A$
$\therefore \Box\Diamond A$

13. $\Box\Diamond A$
$\therefore \Box\Diamond\Box\Diamond A$

14. $\Box(A \supset \Box B)$
$\Diamond A$
$\therefore \Box B$

15. $\Box A$
$\therefore \Box\Box\Box A$

11.1b Exercise: LogiCola KG

Fist appraise intuitively. Then translate into logic (using the letters given) and, assuming S5, prove validity. Also say in which systems the argument is valid: T, B, S4, or S5.

1. It's necessary that if there's a necessary being then "There's a necessary being" (by itself) is necessary.
"There's a necessary being" is logically possible.
∴ "There's a necessary being" is logically necessary. [Use N for "There's a necessary being" or "There's a being that exists of logical necessity"; this being is often identified with God. This argument (which we saw before in §10.3b) is from Charles Hartshorne and St Anselm. Its validity depends on which system of modal logic is correct. Some philosophers defend the argument, often after defending a modal system needed to make it valid. Others argue that the argument is invalid, and so any modal system that would make it valid must be wrong. Still others deny the theological import of the conclusion; they say that a necessary being could be a prime number or the world and needn't be God.]

2. "There's a necessary being" isn't a contingent statement.
"There's a necessary being" is logically possible.
∴ There's a necessary being. [Use N. This version of the Anselm–Hartshorne argument is more clearly valid.] 0254

3. Prove that the first premise of argument 1 is logically equivalent to the first premise of argument 2 by showing that each can be deduced from the other. In which systems does this equivalence hold? 3. It's necessary that if there's a necessary being then "There's a necessary being" (by itself) is necessary.
"There's no necessary being" is logically possible.
∴ There's no necessary being. [Use N. Some object that the first premise of the Anselm–Hartshorne argument just as easily leads to an opposite conclusion.]

4. It's necessary that $2 + 2 = 4$.
It's possible that no language ever existed.
If all necessary truths hold because of language conventions, then "It's necessary that $2 + 2 = 4$" entails "Some language has sometime existed."
∴ Not all necessary truths hold because of language conventions. [Use T, L, and N. This attacks the linguistic theory of logical necessity.]

11.2 Quantified translations

We'll now develop a quantified modal system that combines our quantificational and modal systems. We'll call this our "naïve" system, since it ignores certain problems; later we'll add refinements.[1]

Many quantified modal translations are easy. This pair is tricky:

Everyone could be above average
= ◇(x)Ax
It's possible that everyone is above average
It's possible that, for all x, x is above average

Anyone could be above average
= (x)◇Ax
For all x, it's possible that x is above average

The first is false while the second is true.

Quantified modal logic can express the difference between necessary and contingent properties. Numbers seem to have both kinds of property. The number 8, for example, has the necessary properties of being even and of being one greater than seven; 8 couldn't have lacked these properties. But 8 also has contingent properties, ones it could have lacked, such as being my favorite number and being less than the number of chapters in this book. We can symbolize "necessary property" and "contingent property" as follows: 0255

F is a necessary (essential) property of x
= ☐Fx
x is necessarily F (x has the necessary property of being F)
In all possible worlds, x would be F

F is a contingent (accidental) property of x
= (Fx • ◇~Fx)
x is contingently F (x is F but could have lacked F)
In the actual world x is F; but in some possible world x isn't F.

Humans have mostly contingent properties. Socrates had contingent properties, like having a beard and being a philosopher; these are contingent, because he could (without self-contradiction) have been a clean-shaven non-philosopher. But Socrates also had necessary properties, like being self-identical and not being a square circle; every being has these properties of necessity.

Aristotelian essentialism is the controversial view that there are properties that some beings have of necessity but some other beings totally lack. Plant-

[1] My understanding of quantified modal logic follows Alvin Plantinga's *The Nature of Necessity* (London: Oxford University Press, 1974). For related discussions, see Saul Kripke's *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980) and Kenneth Konyndyk's *Introductory Modal Logic* (Notre Dame, Ind.: Notre Dame Press, 1986).

inga, supporting this view, suggests that Socrates had of necessity these properties that some other beings totally lack: not being a prime number, being snub-nosed in W (a specific possible world), being a person (capable of conscious rational activity), and being identical with Socrates. This last property differs from that of being named "Socrates."

Plantinga explains "necessary property" as follows. Suppose "a" names a being and "F" names a property. Then the entity named by "a" has the property named by "F" necessarily, if and only if the proposition expressed by "a is non-F" is logically impossible. Then to say that Socrates necessarily has the property of not being a prime number is to say that the proposition "Socrates is a prime number" (with the name "Socrates" referring to the person Socrates) is logically impossible. We must use names (like "Socrates") here and not definite descriptions (like "the entity I'm thinking about").

We previously discussed the **box-inside/box-outside** ambiguity. This quantified modal sentence similarly could have either of two meanings:

"All bachelors are necessarily unmarried."

Simple necessity $(x)(Bx \supset \Box Ux)$ All bachelors are inherently unmarriable – in no possible world would anyone marry them.
Conditional necessity $\Box(x)(Bx \supset Ux)$ It's necessarily true that all bachelors are unmarried. (The meaning of "bachelor" makes this true.)

When translating a statement like "All A's are necessarily B's," give both forms. With ambiguous arguments, work out both arguments. As before, fallacies can result from confusing the forms.

Discussions about Aristotelian essentialism frequently involve such modal
0256 ambiguities. This following sentence could have either of two meanings:

"All persons are necessarily persons."

Simple necessity $(x)(Px \supset \Box Px)$ Everyone who in fact is a person has the necessary property of being a person.

Conditional necessity

$\Box(x)(Px \supset Px)$

It's necessary that all persons are persons.

The first is controversial and attributes to each person the necessary property of being a person; the medievals called this *de re* ("of the thing") necessity. If this first form is true, then you couldn't have been a non-person – your existing as a non-person is self-contradictory; this excludes the possibility of your being reincarnated as an unconscious doorknob. In contrast, the second form is trivially true and attributes necessity to the proposition (or saying) "All persons are persons"; the medievals called this *de dicto* ("of the saying") necessity.

11.2a Exercise: LogiCola J (QM & QT)

Translate these English sentences into wffs; translate ambiguous forms both ways.

It's necessary that all mathematicians have the necessary property of being rational.
$\Box(x)(Mx \supset \Box Rx)$ Here the first "$\Box$" symbolizes *de dicto* necessity ("It's necessary that …"), while the second symbolizes *de re* necessity ("have the necessary property of being rational").

1. It's possible for anyone to be unsurpassed in greatness. [Use Ux.]
2. It's possible for everyone to be unsurpassed in greatness.
3. John has the necessary property of being unmarried. [Use Ux and j.]
4. All experts are necessarily smart. [Ex, Sx]
5. Being named "Socrates" is a contingent property of Socrates. [Nx, s]
6. It's necessary that everything is self-identical. [Use "=."]
7. Every entity has the necessary property of being self-identical.
8. John is necessarily sitting. [Sx, j]
9. Everyone observed to be sitting is necessarily sitting. [Ox, Sx]
10. All numbers have the necessary property of being abstract entities. [Nx, Ax]
11. It's necessary that all living beings in this room are persons. [Lx, Px]
12. All living beings in this room have the necessary property of being persons.

13. All living beings in this room have the contingent property of being persons.

14. Any contingent claim could be true. [Cx, Tx]

15. “All contingent claims are true” is possible.

16. All mathematicians are necessarily rational. [Mx, Rx]

17. All mathematicians are contingently two-legged. [Mx, Tx]

18. All mathematical statements that are true are necessarily true. [Mx, Tx] 0257

19. It’s possible that God has the necessary property of being unsurpassed in greatness. [Ux, g]

20. Some being has the necessary property of being unsurpassed in greatness. [Ux]

11.3 Quantified proofs

On our initial naïve approach to quantified modal logic (which has defects), we just use the same quantificational and modal inference rules as before. Here’s a quantified modal proof:

It’s necessary that everything is self-identical.
∴ Every entity has the necessary property of being self-identical.

1 ☐(x)x=x **Valid**
[∴ (x)☐x=x
* 2 ┌ asm: ~(x)☐x=x
* 3 │ ∴ (∃x)~☐x=x {from 2}
* 4 │ ∴ ~☐a=a {from 3}
* 5 │ ∴ ◇~a=a {from 4}
6 │ **W ∴ ~a=a {from 5}**
7 │ W ∴ (x)x=x {from 1}
8 └ **W ∴ a=a {from 7}**
9 ∴ (x)☐x=x {from 2; 6 contradicts 8}

This next modal argument has an ambiguous premise:

All bachelors are necessarily unmarried.
You’re a bachelor.
∴ “You’re unmarried” is logically necessary.

Premise 1 might assert either simple necessity “(x)(Bx ⊃ ☐Ux)” (“All bachelors are *inherently unmarriable*") or conditional necessity ☐(x)(Bx ⊃ Ux)" ("It's necessary that all bachelors are unmarried"). We'll work it out both ways:

Box-inside version (valid but premise 1 is false):

1 (x)(Bx ⊃ ☐Ux) **Valid**
2 Bu
[∴ ☐Uu
3 ┌ **asm: ~☐Uu**
* 4 │ ∴ (Bu ⊃ ☐Uu) {from 1}
5 └ **∴ ☐Uu {from 4 and 2}**
6 ∴ ☐Uu {from 3; 3 contradicts 5}

Box-outside version (invalid):

1 ☐(x)(Bx ⊃ Ux) **Invalid**
2 Bu
[∴ ☐Uu
* 3 asm: ~☐Uu
4 ∴ ◇~Uu {from 3}
5 W ∴ ~Uu {from 4}
6 W ∴ (x)(Bx ⊃ Ux) {from 1}
7 ∴ (x)(Bx ⊃ Ux) {from 1}
* 8 W ∴ (Bu ⊃ Uu) {from 6}
* 9 ∴ (Bu ⊃ Uu) {from 7}
10 W ∴ ~Bu {from 5 and 8}
11 ∴ Uu {from 2 and 9}

	Bu, Uu
W	~Bu, ~Uu

0258 Both versions are flawed; the first has a false premise while the second is invalid. So another proof that you're inherently unmarriable fails! Ambiguous modal arguments often have one interpretation with a false premise and another that's invalid. Such arguments may seem sound until we focus on the ambiguity.

Our refutation has two possible worlds, each with only one entity – you. In the actual world, you're a bachelor and unmarried; in world W, you're not a bachelor and not unmarried. In this galaxy, the premises are true (since in both worlds all bachelors are unmarried – and in the actual world you're a bachelor) but the conclusion is false (since in world W you're not unmarried).

As with relations, applying our proof strategy mechanically sometimes leads into an endless loop. Here we keep getting new letters and worlds, endlessly:

It's possible for anyone to be above average.
∴ It's possible for everyone to be above average.

	1	(x)◇Ax
		[∴ ◇(x)Ax
*	2	asm: ~◇(x)Ax
	3	∴ ☐~(x)Ax {from 2}
*	4	∴ ◇Aa {from 1} **New letter!**
	5	W ∴ Aa {from 4} **New world!**
*	6	W ∴ ~(x)Ax {from 3}
*	7	W ∴ (∃x)~Ax {from 6}
	8	W ∴ ~Ab {from 7} **New letter!**
*	9	∴ ◇Ab {from 1}
	10	WW ∴ Ab {from 9} **New world!**
*	11	WW ∴ ~(x)Ax {from 3}
	12	WW ∴ (∃x)~Ax {from 11}

... and so on endlessly ...

Using ingenuity, we can devise a refutation with two entities and two worlds:

	a, b
W	Aa, ~Ab
WW	Ab, ~Aa

Here each person is above average in some world or other – but in no world is every person above average. For now, we'll assume in our refutations that every world contains the same entities (and at least one such entity).

11.3a Exercise: LogiCola KQ

Say whether valid (and give a proof) or invalid (and give a refutation). 0259

(x)☐Fx
∴ ☐(x)Fx

```
  1  (x)☐Fx    Valid
  [ ∴ ☐(x)Fx
* 2 ┌ asm: ~☐(x)Fx
* 3 │ ∴ ◇~(x)Fx    {from 2}
* 4 │ W ∴ ~(x)Fx    {from 3}
* 5 │ W ∴ (∃x)~Fx    {from 4}
  6 │ W ∴ ~Fa    {from 5}
  7 │ ∴ ☐Fa    {from 1}
  8 └ W ∴ Fa    {from 7}
  9 ∴ ☐(x)Fx    {from 2; 6 contradicts 8}
```

This is called a "Barcan inference," after Ruth Barcan Marcus. It's doubtful that our naïve quantified modal logic gives the right results for arguments like this (see §11.4).

1. (∃x)☐Fx
∴ ☐(∃x)Fx

2. a=b
∴ (☐Fa ⊃ ☐Fb)

3. ∴ ☐(∃x)x=a

4. ∴ (∃x)☐x=a

5. ◇(x)Fx
∴ (x)◇Fx

6. ∴ (x)☐x=x

7. ∴ ☐(x)x=x

8. ☐(x)(Fx ⊃ Gx)
∴ (x)(Fx ⊃ ☐Gx)

9. ◇(∃x)Fx
∴ (∃x)◇Fx

10. (∃x)◇Fx
∴ ◇(∃x)Fx

11. $(\Diamond(x)Fx \supset (x)\Diamond Fx)$
$\therefore ((\exists x){\sim}Fx \supset \Box(\exists x){\sim}Fx)$

12. $\therefore (x)(y)(x{=}y \supset \Box x{=}y)$

13. $\Box(x)(Fx \supset Gx)$
$\Box Fa$
$\therefore \Box Ga$

14. ${\sim}a{=}b$
$\therefore \Box{\sim}a{=}b$

11.3b Exercise: LogiCola KQ

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation). Translate ambiguous English arguments both ways; prove or disprove each symbolization separately.

1. I have a beard.
∴ "Whoever doesn't have a beard isn't me" is a necessary truth. [Use Bx and i. G. E. Moore criticized such reasoning, which he saw as essential to idealistic metaphysics and its claim that every property of a thing is necessary. The conclusion entails that "I have a beard" is logically necessary. Moore would see "Whoever doesn't have a beard isn't me" as only a contingent truth.]

2. "Whoever doesn't have a beard isn't me" is a necessary truth.
∴ "I have a beard" is logically necessary. [Use Bx and i.]

3. Aristotle isn't identical to Plato.
If some being has the property of being necessarily identical to Plato but not all beings have the property of being necessarily identical to Plato, then some beings have necessary properties that other beings lack.
∴ Some beings have necessary properties that other beings lack. [Use a, p, and S (for "Some beings have necessary properties that other beings lack"). This defense of Aristotelian essentialism is essentially from Alvin Plantinga.] 0260

4. All mathematicians are necessarily rational.
Paul is a mathematician.
∴ Paul is necessarily rational. [Mx, Rx, p]

5. Necessarily there exists something unsurpassed in greatness.
∴ There exists something that necessarily is unsurpassed in greatness. [Ux]

6. The number that I'm thinking of isn't necessarily even.
8 = the number that I'm thinking of.
∴ 8 isn't necessarily even.   [Use n, E, and e. Does our naïve quantified modal logic correctly decide whether this argument is valid?]

7. "I'm a thinking being, and there are no material objects" is logically possible.
Every material object has the necessary property of being a material object.
∴ I'm not a material object.   [Use Tx, Mx, and i; from Alvin Plantinga.]

8. All humans are necessarily rational.
All living beings in this room are human.
∴ All living beings in this room are necessarily rational.   [Use Hx, Rx, and Lx; from Aristotle, who was the first logician and the first to combine quantification with modality.]

9. It's not necessary that all cyclists are rational.
Paul is a cyclist.
Paul is rational.
∴ Paul is contingently rational.   [Cx, Rx, p]

10. "Socrates has a pain in his toe but doesn't show pain behavior" is consistent.
It's necessary that everyone who has a pain in his toe is in pain.
∴ "All who are in pain show pain behavior" isn't a necessary truth.   [Use s, Tx for "x has a pain in his toe," Bx for "x shows pain behavior," and Px for "x is in pain." This attacks a behaviorist analysis of the concept of "pain."]

11. If Q (the question "Why is there something and not nothing?") is a meaningful question, then it's possible that there's an answer to Q.
Necessarily, every answer to Q refers to an existent that explains the existence of other things.
Necessarily, nothing that refers to an existent that explains the existence of other things is an answer to Q.
∴ Q isn't a meaningful question.   [M, Ax, Rx]

12. The number of apostles is 12.
12 is necessarily greater than 8.
∴ The number of apostles is necessarily greater than 8.   [Use n, t, e, and Gxy. Does our naïve system correctly decide whether this argument is valid?]

13. All (well-formed) cyclists are necessarily two-legged.
Paul is a (well-formed) cyclist.
∴ Paul is necessarily two-legged.   [Cx, Tx, p] 0261

14. Something exists in the understanding than which nothing could be greater. (In other words, there's some x such that x exists in the understanding and it's not possible that there be something greater that x.)
Anything that exists in reality is greater than anything that doesn't exist in reality.
Socrates exists in reality.
∴ Something exists in reality than which nothing could be greater. (In other words, there's some x such that x exists in reality and it's not possible that there be something greater than x.) [Use Ux for "x exists in the understanding," Rx for "x exists in reality," Gxy for "x is greater than y," and s for "Socrates." Use a universe of discourse of *possible beings* – including fictional beings like Santa Claus in addition to actual beings. (Is this legitimate?) This is a form of St Anselm's first ontological argument for the existence of God.]

15. "Someone is unsurpassably great" is logically possible.
"Everyone who is unsurpassably great is, in every possible world, omnipotent, omniscient, and morally perfect" is necessarily true.
∴ Someone is omnipotent, omniscient, and morally perfect. [Use Ux and Ox. This is a simplified form of Alvin Plantinga's ontological argument for the existence of God. Plantinga regards the second premise as true by definition; he sees the first premise as controversial but reasonable.]

16. Anything could cease to exist.
∴ Everything could cease to exist. [Use Cx for "x ceases to exist." Some see Aquinas's third argument for the existence of God as requiring this inference.]

11.4 A sophisticated system

Our naïve quantified modal logic has defects. Dealing with these will push us to question established logical and metaphysical ideas.

First, our system mishandles definite descriptions (terms of the form "the so and so"). We've been translating definite descriptions using small letters, as in the following example:

The number I'm thinking of is necessarily odd = ☐On

But this English sentence is ambiguous; it could mean either of two things (where "Tx" means "I'm thinking of number x"):

$(\exists x)((Tx \bullet \sim(\exists y)(\sim x{=}y \bullet Ty)) \bullet \Box Ox)$
I'm thinking of just one number, and it has the necessary property of being odd.

$\Box(\exists x)((Tx \bullet \sim(\exists y)(\sim x{=}y \bullet Ty)) \bullet Ox)$
This is necessary: "I'm thinking of just one number and it's odd."

The first form (box inside) might be true – if, for example, the number 7 has the necessary property of being odd and I'm thinking of just the number 7. The 0262 second form (box outside) is definitely false, since it's possible that I'm thinking of no number, or more than one number, or an even number.

So our naïve way to translate "the so and so" is ambiguous. To fix this problem, our sophisticated system will require that we symbolize "the so and so" using Russell's "there is just one …" analysis (§9.6) – as in the above boxes. This analysis also blocks the proof of invalid arguments like this one:

8 is the number I'm thinking of.
It's necessary that 8 is 8.
∴ It's necessary that 8 is the number I'm thinking of.

$e{=}n$
$\Box e{=}e$
$\therefore \Box e{=}n$

This is invalid – since it may be only contingently true that 8 is the number I'm thinking of. The argument is provable in naïve quantified modal logic, since the conclusion follows from the premises by the substitute-equals rule (§9.2). Our sophisticated system avoids this by requiring the longer analysis of "the number I'm thinking of." So "8 is the number I'm thinking of" gets changed into "I'm thinking of just one number and it is 8" – and the above argument becomes this:

I'm thinking of just one number and it is 8.
It's necessary that 8 is 8.
∴ This is necessary: "I'm thinking of just one number and it is 8."

$(\exists x)((Tx \bullet \sim(\exists y)(\sim x{=}y \bullet Ty)) \bullet x{=}e)$ **Invalid**
$\Box e{=}e$
$\therefore \Box(\exists x)((Tx \bullet \sim(\exists y)(\sim x{=}y \bullet Ty)) \bullet x{=}e)$

So translated, the argument becomes invalid and not provable.

The second problem is that our naïve system assumes that the same entities exist in all possible worlds. This leads to implausible results; for example, it makes Gensler (and everyone else) into a logically necessary being:

∴ In every possible world, there exists a being who is Gensler.

[∴ ☐(∃x)x=g **Valid ???**
* 1 ┌ asm: ~☐(∃x)x=g
* 2 │ ∴ ◇~(∃x)x=g {from 1}
* 3 │ W ∴ ~(∃x)x=g {from 2}
4 │ W ∴ (x)~x=g {from 3}
5 │ W **∴ ~g=g {from 4} ???**
6 └ W **∴ g=g {self-identity}**
7 ∴ ☐(∃x)x=g {from 1; 5 contradicts 6}

But Gensler isn't a logically necessary being; there are impoverished possible worlds without me. So something is wrong here.

There are two ways out of the problem. One way changes how we take "(∃x)." The provable "☐(∃x)x=g" is false if we take "(∃x)" to mean "for some *existing being* x." But we might take "(∃x)" to mean "for some *possible being* x"; then "☐(∃x)x=g" would mean the more plausible: "In every possible world, there's a *possible being* who is Gensler." Perhaps there's a possible being Gensler in every 0263 world; in some of these worlds Gensler exists, and in others he doesn't. This view would need an existence predicate "Ex" to distinguish between possible beings that exist and those that don't; we could then use "(∃x)~Ex" to say that there are possible beings that don't exist. This view is paradoxical, since it posits non-existent beings.

Alvin Plantinga defends an opposing view, which he calls "actualism." *Actualism* holds that to be a being and to exist is the same thing; there neither are nor could have been non-existent beings. Of course there could have been beings other than those that now exist. But this doesn't mean that there *now* are beings that don't exist. Actualism denies the latter claim.

Since I favor actualism, I'll avoid non-existent beings and continue to take "(∃x)" to mean "for some existing being." On this reading, "☐(∃x)x=g" means "It's necessary that there's an existing being who is Gensler." This is false, since I might not have existed. So we must reject some line of the above proof.

The faulty line seems to be 5 (and its derivation from 4):

In W, every existing being is distinct from Gensler.
∴ In W, Gensler is distinct from Gensler.

4 W ∴ (x)~x=g
5 W ∴ ~g=g {from 4}

This inference shouldn't be valid – unless we presuppose the additional premise "W ∴ (∃x)x=g" – that Gensler is an existing being in world W.

Rejecting line 5 requires moving to a *free logic* – one free of the assumption that individual constants like "g" always refer to existing beings. Recall our drop-universal rule DU of §8.2:

Drop universal DU (x)Fx → Fa, use any constant
Every existing being is F. ∴ a is F.

Suppose that every existing being is F; "a" might not denote an existing being, and so "a is F" might not be true. So we need to modify the rule to require the premise that "a" denotes an existing being:

Drop universal DU* (x)Fx, (∃x)x=a → Fa, use any constant
Every existing being is F. a is an existing being. ∴ a is F.

Here we symbolize "a is an existing being" by "(∃x)x=a" ("For some existing being x, x is identical to a"). With this change, "☐(∃x)x=g" ("Gensler is a necessary being") is no longer provable.

If we weaken DU, we need to strengthen our drop-existential rule DE:

Drop existential DE* (∃x)Fx → Fa, (∃x)x=a, use a *new constant*
Some existing being is F. ∴ a is F. ∴ a is an existing being.

0264 When we drop an existential using DE*, we get an existence claim (like "(∃x)x=a") that we can use in dropping universals with DU*. The resulting system can prove almost everything we could prove before – except that proofs are now longer. The main effect is to block a few proofs; we can no longer prove that Gensler exists in all possible worlds.

Our free-logic system also blocks the proof of this Barcan inference:

Every existing being has the necessary property of being F.
∴ In every possible world, every existing being is F.

```
   1  (x)☐Fx    Invalid
   [ ∴ ☐(x)Fx
 * 2  asm: ~☐(x)Fx
 * 3  ∴ ◇~(x)Fx    {from 2}
 * 4  W ∴ ~(x)Fx    {from 3}
 * 5  W ∴ (∃x)~Fx  {from 4}
   6  W ∴ ~Fa    {from 5}
   7  W ∴ (∃x)x=a  {from 5}
```

	b exists, a doesn't Fb, ~Fa
W	a & b exist Fb, ~Fa

Our new rule for dropping "(∃x)" tells us that "a" denotes an existing being in world W (line 7). But we don't know if "a" denotes an existing being in the actual world; so we can't conclude "☐Fa" from "(x)☐Fx" in line 1. With our naïve system, we could conclude "☐Fa" – and then put "Fa" in world W to contradict line 6; but now the line is blocked, and the proof fails.

While we don't automatically get a refutation, we can invent one on our own. Our refutation lists which entities exist in which worlds; it uses "a exists" for "(∃x)x=a." Here "Every existing being has the necessary property of being F" is true – since entity b is the only existing being and in every world it is F. But "In every possible world, every existing being is F" is false – since in world W there is an existing being, a, that isn't F.

Here's another objection to the argument. Suppose only abstract objects (numbers, sets, etc.) existed and all these had the necessary property of being abstract. Then "Every existing being has the necessary property of being abstract" would be true. But "In every possible world, every existing being is abstract" could still be false – if other possible worlds had concrete entities.[1]

[1] Or suppose God created nothing and all uncreated beings had the necessary property of being uncreated. Then "Every existing being has the necessary property of being uncreated"

Our new approach lets different worlds have different existing entities. Gensler might exist in one world but not another. We shouldn't picture existing in different worlds as spooky; it's just a way of talking about different possibilities. I might not have existed. We can tell consistent stories where my parents didn't meet and where I never came into existence. If the stories had been true, then I wouldn't have existed. So I don't exist in these stories (although I might exist in other stories). Existing in a possible world is much like existing in a story; a "possible world" is a technical analogue of a "consistent story." "I exist in world W" just means "If world W had been actual, then I would have existed." 0265

We also could allow possible worlds with no entities. In such worlds, all wffs starting with existential quantifiers are false and all those starting with universal quantifiers are true.

Should we allow this as a possible world when we do our refutations?

W	a doesn't exist, Fa

It seems incoherent to claim that "a has property F" is true while a doesn't exist. It seems that only existing beings have positive properties; in a consistent story where Gensler doesn't exist, Gensler couldn't be a logician or a backpacker. So if "a exists" isn't true in a possible world, then "a has property F" isn't true in that world either. We can put this idea into an inference rule PE*:

Property existence PE*
$Fa \rightarrow (\exists x)x{=}a$
a has property F. ∴ a is an existing being.

Rule PE* holds regardless of what capital letter replaces "F," what constant replaces "a," and what variable replaces "x." By PE*, "Descartes thinks" entails "Descartes exists." Conversely, the falsity of "Descartes exists" entails the falsity of "Descartes thinks." Rule PE* expresses that it's a necessary truth that only existing objects have properties. Plantinga calls this view "serious actualism"; actualists who reject PE* are deemed frivolous.

The first example below isn't a correct instance of PE* (since the wff substituted for "Fa" in PE* can't contain "∼"), but the second is:

would be true. But "In every possible world, every existing being is uncreated" could still be false – since there could have been possible worlds with created beings.

This one is wrong:

$\underline{\sim Fa\qquad}$
$\therefore (\exists x)x=a$

a isn't F
∴ a exists

This one is right:

$\underline{Fa\qquad\quad}$
$\therefore (\exists x)x=a$

a is F
∴ a exists

This point is confusing because "a isn't F" in English can have two different senses. "Descartes doesn't think" could mean either of these:

Descartes is an existing being who doesn't think
= $(\exists x)(x=d \bullet \sim Td)$

It's false that Descartes is an existing being who thinks
= $\sim(\exists x)(x=d \bullet Td)$

The first form is *de re* (about the thing); it affirms the property of being a non-thinker of the entity Descartes. Taken this first way, "Descartes doesn't think" entails "Descartes exists." The second form is *de dicto* (about the saying); it denies the statement "Descartes thinks" (which may be false either because Descartes is a non-thinking entity or because Descartes doesn't exist). Taken this second way, "Descartes doesn't think" doesn't entail "Descartes exists."

One might object to PE* on the grounds that Santa Claus has properties (such as being fat) but doesn't exist. But various stories predicate conflicting properties 0266 to Santa; they differ, for example, on which day he delivers presents. Does Santa have contradictory properties? Or is one Santa story uniquely "true"? What would that mean? When we say "Santa is fat," we mean that in such and such a story (or possible world) there's a being called Santa who is fat. We shouldn't think of Santa as a non-existing being in our actual world who has properties such as being fat. Rather, what exists in our actual world is stories about there being someone with certain properties – and children who may believe these stories. So Santa needn't make us give up PE*.

We need to modify our current definition of "necessary property":

F is a necessary property of a
= $\Box$Fa
In all possible worlds, a is F

Let's grant that Socrates has properties only in worlds where he exists – and that there are worlds where he doesn't exist. Then there are worlds where Socrates has no properties – and so there aren't any properties that Socrates has in all worlds. By our definition, Socrates would have no necessary properties.

Socrates still might have some necessary *combinations* of properties. Perhaps it's true in all worlds that if Socrates exists then Socrates is a person. This suggests a more refined definition of "necessary property":

F is a necessary property of a
= $\Box$((∃x)x=a ⊃ Fa)
It's necessary that if a exists then a is F
In all possible worlds where a exists, a is F

This reflects better what philosophers mean when they speak of *necessary properties*. It also lets us claim that Socrates has the necessary property of being a person. This would mean that Socrates is a person in every possible world where he exists; equivalently, in no possible world does Socrates exist as anything other than a person. Here's an analogous definition of "contingent property":

F is a contingent property of a
= (Fa • $\Diamond$((∃x)x=a • ~Fa))
a is F; but in some possible world where a exists, a isn't F

This section sketched a sophisticated quantified modal logic. Its refinements overcome some problems but also make the system harder to use. We seldom need the refinements. So we'll keep the naïve system of earlier sections as our "official system" and build on it in later chapters. But we'll be aware that this system is oversimplified in some ways. If our naïve system gives questionable results, we can appeal to the sophisticated system to clear things up.

12 Deontic and Imperative Logic

Imperative logic studies arguments with imperatives, like "Don't do this." **Deontic logic** studies arguments whose validity depends on "ought," "permissible," and similar notions. We'll take imperative logic first and then build deontic logic on it.[1]

12.1 Imperative translations

Imperative logic builds on previous systems and adds two ways to form wffs:

> 1. Any underlined capital letter is a wff.
> 2. The result of writing a capital letter and then one or more small letters, one small letter of which is underlined, is a wff.

<u>Underlining</u> (combined with <u>**bolding**</u> in this e-book version) turns indicatives into imperatives:

Indicative (You're doing A)
A Au
Imperative (Do A)
$\underline{\mathbf{A}}$ $A\underline{\mathbf{u}}$

Here are some further translations:

[1] I'll mostly follow Hector-Neri Castañeda's approach. See his "Imperative reasonings," *Philosophy and Phenomenological Research* 21 (1960): pp. 21–49; "Outline of a theory on the general logical structure of the language of action," *Theoria* 26 (1960): pp. 151–82; "Actions, imperatives, and obligations," *Proceedings of the Aristotelian Society* 68 (1967–68): pp. 25–48; and "On the semantics of the ought-to-do," *Synthese* 21 (1970): pp. 448–68.

Don’t do A = $\sim\underline{\mathbf{A}}$
Do A and B = $(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$
Do A or B = $(\underline{\mathbf{A}} \vee \underline{\mathbf{B}})$
Don’t do either A or B = $\sim(\underline{\mathbf{A}} \vee \underline{\mathbf{B}})$ 0268

Don’t combine doing A with doing B
= $\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$
Don’t both do A and do B

Don’t combine doing A with not doing B
= $\sim(\underline{\mathbf{A}} \bullet \sim\underline{\mathbf{B}})$
Don’t do A without doing B

Underline imperative parts but not factual ones:

You’re doing A and you’re doing B = $(A \bullet B)$
You’re doing A, but do B = $(A \bullet \underline{\mathbf{B}})$
Do A and B = $(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

If you’re doing A, then you’re doing B
= $(A \supset B)$

If you (in fact) are doing A, then do B
= $(A \supset \underline{\mathbf{B}})$

Do A, only if you (in fact) are doing B
= $(\underline{\mathbf{A}} \supset B)$

Since English can’t put an imperative after “if,” we can’t read “$(\underline{\mathbf{A}} \supset B)$” as “If do A, then you’re doing B.” But we can read it as the equivalent “Do A, only if you’re doing B.” This means the same as “$(\sim B \supset \sim\underline{\mathbf{A}})$”: “If you aren’t doing B, then don’t do A.”

There’s a subtle difference between these two:

If you (in fact) are doing A, then don’t do B
= $(A \supset \sim\underline{\mathbf{B}})$

Don’t combine doing A with doing B
= $\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

“A” is underlined in the second but not the first; otherwise, the two wffs would be equivalent. The if-then “$(A \supset \sim\underline{\mathbf{B}})$” says that if A is done then you aren't to do B. But the don't-combine "$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$" just forbids a combination: doing A and B together. If you're doing A, it doesn't follow that you aren't to do B; maybe you should do B and stop doing A. We'll see more on this distinction later.

These examples underline the letter for the agent:

X, do (or be) A = $A\underline{x}$
X, do A to Y = $A\underline{x}y$

These use quantifiers:

Everyone does A = $(x)Ax$
Let everyone do A = $(x)A\underline{x}$

Let everyone who (in fact) is doing A do B
= $(x)(Ax \supset B\underline{x})$

Let someone who (in fact) is doing A do B
= $(\exists x)(Ax \bullet B\underline{x})$

Let someone both do A and do B
= $(\exists x)(A\underline{x} \bullet B\underline{x})$

Notice which letters are underlined. 0269

12.1a Exercise: LogiCola L (IM & IT)

Translate these English sentences into wffs; take each "you" as a singular "you."

If the cocoa is about to boil, remove it from the heat
$(B \supset \underline{R})$ Our sentence also could translate as "$(B \supset R\underline{u})$" or "$(Bc \supset R\underline{u}c)$."

1. Leave or shut up. [Use L and S.]
2. If you don't leave, then shut up.
3. Do A, only if you want to do A. [Use A and W.]
4. Do A, only if you want to do A. [This time use Au and Wu.]
5. Don't combine accelerating with braking.

6. If you accelerate, then don’t brake.

7. If you brake, then don’t accelerate.

8. If you believe that you ought to do A, then do A. [Use A for “You do A” and B for “You believe that you ought to do A.”]

9. Don’t combine believing that you ought to do A with not doing A.

10. If everyone does A, then do A yourself.

11. If you have a headache, then take aspirin. [Hx, Ax, u]

12. Let everyone who has a headache take aspirin.

13. Gensler, rob Jones. [Rxy, g, j]

14. If Jones hits you, then hit Jones. [Hxy, j, u]

15. If you believe that A is wrong, then don’t do A. [Use A for “You do A” and B for “You believe that A is wrong.”]

16. If you do A, then don’t believe that A is wrong.

17. Don’t combine believing that A is wrong with doing A.

18. Would that someone be sick and also be well. [Sx, Wx]

19. Would that someone who is sick be well.

20. Would that someone be sick who is well.

12.2 Imperative proofs

Imperative proofs work much like indicative ones and require no new inference rules. But we must treat “A” and “$\underline{\mathbf{A}}$” as different wffs. “A” and “$\sim\underline{\mathbf{A}}$” aren’t contradictories; it’s consistent to say “You’re now doing A, but don’t.”

Here’s an imperative argument that follows an I-rule inference:

If you’re accelerating, then don’t brake.
You’re accelerating.
∴ Don’t brake.

$(\mathrm{A} \supset \sim\underline{\mathbf{B}})$ **Valid**
A
∴ $\sim\underline{\mathbf{B}}$

While this seems valid, there’s a problem with calling it “valid.” We earlier defined “valid” using “true” and “false” (§1.2): an argument is *valid* if it would
0270 be contradictory to have the premises all *true* and conclusion *false*. But “Don’t brake” and other imperatives aren’t true or false. So how can the valid

/invalid distinction apply to imperative arguments?

We need a broader definition of "valid" that applies equally to indicative and imperative arguments. This one (which avoids "true" and "false") does the job:

An argument is *valid* if the conjunction of its premises with its conclusion's contradictory is inconsistent.

To say that our argument is *valid* means that this combination is inconsistent:

"If you're accelerating, then don't brake; you're accelerating; brake."

The combination *is* inconsistent. So our argument is *valid* in this new sense.[1]

This next argument uses a *don't-combine* premise, which makes it invalid:

Don't combine accelerating with braking.
You're accelerating.
∴ Don't brake.

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ **Invalid**
A
∴ $\sim\underline{\mathbf{B}}$

The first premise forbids us to accelerate and brake together. Suppose we're accelerating. It doesn't follow that we shouldn't brake; maybe, to avoid hitting a car, we should brake and stop accelerating. So the argument is invalid. It's consistent to conjoin the premises with the contradictory of the conclusion:

Don't combine accelerating with braking – never do both together; you in fact are accelerating right now; but you'll hit a car unless you slow down; so stop accelerating right away – and brake immediately.

Here it makes good consistent sense to endorse the premises while also adding the denial of the conclusion ("Brake").

We'd work out the symbolic argument this way (being careful to treat "A" and "$\underline{\mathbf{A}}$" as different wffs, almost as if they were different letters):

[1] We could equivalently define a *valid argument* as one in which every set of imperatives and indicatives that's consistent with the premises also is consistent with the conclusion.

* 1 $\sim(\underline{\mathbf{A}}^{0} \bullet \underline{\mathbf{B}}^{1}) = 1$ **Invalid**
2 $A^{1} = 1$
[$\therefore \sim\underline{\mathbf{B}}^{1} = 0$
3 asm: $\underline{\mathbf{B}}$
4 $\therefore \sim\underline{\mathbf{A}}$ {from 1 and 3}

A, $\sim\underline{\mathbf{A}}$, $\underline{\mathbf{B}}$

On our refutation:

$A = 1$
$\underline{\mathbf{A}} = 0$
$\underline{\mathbf{B}} = 1$

We quickly get a refutation – a set of assignments of 1 and 0 to the letters that make the premises 1 but conclusion 0. Our refutation says this: 0271

You're accelerating; don't accelerate; instead, brake.

But our refutation assigns *false* to the imperative "Accelerate" – even though imperatives aren't true or false. So what does "$\underline{\mathbf{A}} = 0$" mean?

We can generically read "1" as "correct" and "0" as "incorrect." Applied to indicatives, these mean "true" or "false." Applied to imperatives, these mean that the prescribed action is "correct" or "incorrect" relative to some standard that divides actions prescribed by the imperative letters into *correct* and *incorrect* actions. The standard could be of different sorts, based on things like morality, law, or traffic safety; generally we won't specify the standard.

Suppose we have a propositional-logic argument with imperative letters added. The argument is *valid* if and only if, relative to every assignment of "1" or "0" to the indicative and imperative letters, if the premises are "1," then so is the conclusion. Equivalently, the argument is *valid* if and only if, relative to any possible facts and any possible consistent standards for correct actions, if all the premises are correct then so is the conclusion.

So our refutation amounts to this: we imagine certain facts being true/false and certain actions being correct/incorrect:

$A = 1$ "You're accelerating" is true.
$\underline{\mathbf{A}} = 0$ Accelerating is incorrect.
$\underline{\mathbf{B}} = 1$ Braking is correct.

Our argument could have all the premises correct but not the conclusion.

Compare the two imperative arguments that we've considered:

If you're accelerating, then don't brake.

You're accelerating.
∴ Don't brake.

$(A \supset \sim\underline{\mathbf{B}})$ **Valid**
A
$\therefore \sim\underline{\mathbf{B}}$

Don't combine accelerating with braking.
You're accelerating.
∴ Don't brake.

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ **Invalid**
A
$\therefore \sim\underline{\mathbf{B}}$

Both arguments are the same, except that the first uses an if-then "$(A \supset \sim\underline{\mathbf{B}})$," while the second uses a don't-combine "$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$." Since one argument is valid and the other isn't, the two wffs aren't equivalent.

Imagine that you find yourself accelerating and braking, thus wearing down your brakes and wasting energy. Then you violate all three of these imperatives:

$(A \supset \sim\underline{\mathbf{B}})$ = If you're accelerating, then don't brake

$(B \supset \sim\underline{\mathbf{A}})$ = If you're braking, then don't accelerate

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ = Don't combine accelerating with braking

The three differ on what to do next. The first tells you not to brake. The second tells you not to accelerate. But the third leaves it open whether you're to stop 0272 accelerating or stop braking. Maybe you need to brake (and stop accelerating) to avoid hitting another car; or maybe you need to accelerate (and stop braking) to pass another car. The don't-combine form doesn't tell a person in this forbidden combination exactly what to do.

Consistency imperatives need the don't-combine form. Suppose that you're inconsistent if you combine *doing A* with *doing B*. Then:

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ = Don't combine doing A with doing B.

This forbids a combination but doesn't say exactly what to do. Suppose that you're inconsistently doing A and B together. From this we can't conclude which you are to change; both of these are invalid:

Don't combine doing A with doing B.

You're doing A.
∴ Don't do B.

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ **Invalid**
A
∴ $\sim\underline{\mathbf{B}}$

Don't combine doing A with doing B.
You're doing B.
∴ Don't do A.

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ **Invalid**
B
∴ $\sim\underline{\mathbf{A}}$

These inference forms are wrong, even though they may seem correct. Together they'd tell you to give up *both* A and B. But all you need to do is give up *one* of these, A or B. The "$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$" form is logically equivalent to "$(\sim\underline{\mathbf{A}} \vee \sim\underline{\mathbf{B}})$," which means "Either don't do A or don't do B."

Suppose that *acting to do this* is somehow inconsistent with *believing that this is wrong*. Here's the corresponding consistency imperative:

$\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ = Don't combine acting to do this with believing that this is wrong

This combination always has a faulty element. If your act is correct, then your belief is wrong; if your belief is correct, then your act is wrong. If you combine this act with this belief, then your act clashes with your belief. How should you regain consistency? This depends on the situation – since either of the two could be faulty; so sometimes it's better to change your act and sometimes it's better to change your belief.[1] The don't-combine form forbids an inconsistency, but it correctly doesn't tell a person in this forbidden combination exactly what to do. For this reason, it's important to express consistency imperatives as pure don't-combine imperatives instead of as mixed if-then imperatives like these:

$(B \supset \sim\underline{\mathbf{A}})$ = If you believe that this is wrong, then don't act to do this

$(A \supset \sim\underline{\mathbf{B}})$ = If you act to do this, then don't believe that this is wrong 0273

The first wrongly assumes that your belief has to be correct in such conflict cases, while the second wrongly assumes that your act has to be correct.

[1] Maybe your act is fine but your belief is faulty; for example, you treat dark-skinned people fairly but believe that this is wrong. More typically, your belief is fine but your act is faulty.

Since either can be faulty, both if-then imperatives can give bad advice. So it's better to express consistency imperatives as don't-combine forms, like "∼(**A** • **B**)."

Before leaving this section, let me point out problems with two alternative ways to understand imperative logic. Consider this argument:

If you get 100 percent, then celebrate.
Get 100 percent.
∴ Celebrate.

(G ⊃ **C**) **Invalid**
G
∴ **C**

G, ∼G, ∼**C**

This is intuitively invalid. Don't celebrate yet – maybe you'll flunk. To derive the conclusion, we need, not an imperative second premise, but rather a factual one saying that you *did* get 100 percent.

Two common ways to understand imperative logic would wrongly judge this argument to be valid. The *obedience view* says that an imperative argument is valid just if doing what the premises prescribe necessarily involves doing what the conclusion prescribes. This is fulfilled in the present argument; if you do what both premises say, you'll get 100 percent and celebrate. So the obedience view says that our argument is valid. So the obedience view is wrong.

The *threat view* analyzes the imperative "Do A" as "Either you will do A or else S will happen" – where sanction "S" is some unspecified bad thing. So "**A**" is taken to mean "(A ∨ S)." But if we replace "**C**" with "(C ∨ S)" in our argument and "**G**" with "(G ∨ S)," then our argument becomes valid. So the threat view says that our argument is valid. So the threat view is wrong.

12.2a Exercise: LogiCola MI

Say whether valid (and give a proof) or invalid (and give a refutation).

(A ⊃ ∼**B**) (∼A ⊃ ∼**C**) ∴ ∼(**B** • **C**)

* 1 $(A \supset \sim\underline{\mathbf{B}})$ **Valid**
* 2 $(\sim A \supset \sim\underline{\mathbf{C}})$
[$\therefore \sim(\underline{\mathbf{B}} \bullet \underline{\mathbf{C}})$
* 3 asm: $(\underline{\mathbf{B}} \bullet \underline{\mathbf{C}})$
4 $\therefore \underline{\mathbf{B}}$ {from 3}
5 $\therefore \underline{\mathbf{C}}$ {from 3}
6 $\therefore \sim A$ {from 1 and 4}
7 $\therefore A$ {from 2 and 5}
8 $\therefore \sim(\underline{\mathbf{B}} \bullet \underline{\mathbf{C}})$ {from 3; 6 contradicts 7}

1. $\sim\underline{\mathbf{A}}$
$\therefore \sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

2. $\sim(\underline{\mathbf{A}} \bullet \sim\underline{\mathbf{B}})$
$\therefore (A \supset \underline{\mathbf{B}})$

3. $(A \supset \underline{\mathbf{B}})$
$\therefore (\sim B \supset \sim\underline{\mathbf{A}})$

4. $(A \supset \underline{\mathbf{B}})$
$\therefore \sim(\underline{\mathbf{A}} \bullet \sim\underline{\mathbf{B}})$

5. $\sim\Diamond(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$
$\sim(\underline{\mathbf{C}} \bullet \sim\underline{\mathbf{A}})$
$\therefore \sim(\underline{\mathbf{C}} \bullet \underline{\mathbf{B}})$ 0274

6. $(x)(Fx \supset G\underline{\mathbf{x}})$
$F\underline{\mathbf{a}}$
$\therefore G\underline{\mathbf{a}}$

7. $(x)\sim(F\underline{\mathbf{x}} \bullet G\underline{\mathbf{x}})$
$(x)(Hx \supset F\underline{\mathbf{x}})$
$\therefore (x)(G\underline{\mathbf{x}} \supset \sim Hx)$

8. $(x)(Fx \supset G\underline{\mathbf{x}})$
$(x)(Gx \supset H\underline{\mathbf{x}})$
$\therefore (x)(Fx \supset H\underline{\mathbf{x}})$

9. $(\sim\underline{\mathbf{A}} \vee \sim\underline{\mathbf{B}})$
$\therefore \sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

10. $\sim(\underline{\mathbf{A}} \bullet \sim\underline{\mathbf{B}})$
$\therefore (\sim\underline{\mathbf{A}} \vee \underline{\mathbf{B}})$

12.2b Exercise: LogiCola MI

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (and give a refutation).

1. Make chicken for dinner or make eggplant for dinner.
Peter is a vegetarian.
If Peter is a vegetarian, then don't make chicken for dinner.
∴ Make eggplant for dinner.   [Use C, E, and V. This one is from Peter Singer.]

2. Don't eat cake.
If you don't eat cake, then give yourself a gold star.
∴ Give yourself a gold star.   [Use E and G.]

3. If this is greasy food, then don't eat this.
This is greasy food.
∴ Don't eat this.   [Use G and E; from Aristotle, except that he saw the conclusion of an imperative argument as an action: since you accept the premises, you don't eat the thing. I'd prefer to say that if you accept these premises and are consistent, then you won't eat the thing.]

4. Don't both drive and watch the scenery.
Drive.
∴ Don't watch the scenery.   [D, W]

5. If you believe that you ought to commit mass murder, then commit mass murder.
You believe that you ought to commit mass murder.
∴ Commit mass murder.   [Use B and C. Suppose we take "Follow your conscience" to mean "If you believe that you ought to do A, then do A." Then this principle can tell us to do evil things. Would the corresponding don't-combine form also tell us to do evil things? See the next example.]

6. Don't combine believing that you ought to commit mass murder with not committing mass murder.
You believe that you ought to commit mass murder.
∴ Commit mass murder.   [B, C]

7. Don't combine having this end with not taking this means.
Don't take this means.
∴ Don't have this end.   [E, M] 0275

8. Lie to your friend only if you want people to lie to you under such circumstances.
You don't want people to lie to you under such circumstances.
∴ Don't lie to your friend. [Use L and W. Premise 1 is based on a simplified version of Immanuel Kant's formula of universal law; we'll see a more sophisticated version in Chapter 14.]

9. Studying is needed to become a teacher.
"Become a teacher" entails "Do what's needed to become a teacher."
"Do what's needed to become a teacher" entails "If studying is needed to become a teacher, then study."
∴ Either study or don't become a teacher. [Use N for "Studying is needed to become a teacher," B for "You become a teacher," D for "You do what's needed to become a teacher," and S for "You study." This example shows that we can deduce complex ends-means imperatives from purely descriptive premises.]

10. Winn Dixie is the largest grocery store in Big Pine Key.
∴ Either go to Winn Dixie or don't go to the largest grocery store in Big Pine Key. [w, l, Gxy, u]

11. Drink something available.
Only juice and soda are available.
∴ Drink some juice or soda. [Dxy, u, Ax, Jx, Sx]

12. If the cocoa is about to boil, remove it from the heat.
If the cocoa is steaming, it's about to boil.
∴ If the cocoa is steaming, remove it from the heat. [B, R, S]

13. Don't shift.
∴ Don't combine shifting with not pedaling. [S, P]

14. If he's in the street, wear your gun.
Don't wear your gun.
∴ He isn't in the street. [Use S and G. This imperative argument, from Hector-Neri Castañeda, has a factual conclusion; calling it "valid" means that it's inconsistent to conjoin the premises with the denial of the conclusion.]

15. If you take logic, then you'll make logic mistakes.
Take logic.
∴ Make logic mistakes. [T, M]

16. Get a soda.
If you get a soda, then pay a dollar.
∴ Pay a dollar. [G, P]

17. ∴ Either do A or don't do A. [This (vacuous) imperative tautology is analogous to the logical truth "You're doing A or you aren't doing A."]

18. Don't combine believing that A is wrong with doing A.
∴ Either don't believe that A is wrong, or don't do A. [B, A] 0276

19. Mail this letter.
∴ Mail this letter or burn it. [Use M and B. This one was used to try to discredit imperative logic. The argument is valid, since this is inconsistent: "Mail this letter; don't either mail this letter or burn it." Note that "Mail this letter or burn it" doesn't entail "You may burn it"; it's consistent to follow "Mail this letter or burn it" with "Don't burn it."]

20. Let every incumbent who will be honest be endorsed.
∴ Let every incumbent who won't be endorsed not be honest. [Use Hx, Ex, and the universe of discourse of incumbents.]

12.3 Deontic translations

Deontic logic adds two operators: "O" (for "ought") and "R" (for "all right" or "permissible"); these attach to imperatives to form deontic wffs:

O**<u>A</u>**	=	It's obligatory that A
OA**<u>u</u>**	=	You ought to do A
R**<u>A</u>**	=	It's permissible that A
RA**<u>u</u>**	=	It's all right for you to do A

"O"/"☐" (moral/logical necessity) are somewhat analogous, as are "R"/"◇" (moral/logical possibility).

"Ought" here is intended in the all-things-considered, normative sense that we often use in discussing moral issues. This sense of "ought" differs from at least two other senses that may follow different logical patterns:

- *Prima facie* senses of "ought" (which give a moral consideration that may be overridden in a given context): "Insofar as I promised to go with you to the movies, I ought to do this [*prima facie* duty]; but insofar as my wife needs me to drive her to the hospital, I ought to do this instead [*prima facie* duty]. Since my duty to my wife is more weighty, in the final analysis I ought to drive my wife to the hospital [all-things-considered duty]."
- Descriptive senses of "ought" (which state what's required by conventional

social rules but needn't express one's own positive or negative evaluation): "You ought [by company regulations] to wear a tie to the office."

I'll be concerned with logical connections between ought judgments, where "ought" is taken in this all-things-considered, normative sense.[1] I'll mostly avoid metaethical issues, like how to further analyze "ought," how to justify ethical principles, and whether moral judgments are objectively true or false. While my 0277 explanations sometimes assume that ought judgments are true or false, what I say could be rephrased to avoid this assumption.[2]

Here are some further translations:

Act A is obligatory (required, a duty)
= $O\underline{\mathbf{A}}$

Act A is all right (right, permissible, OK)
= $R\underline{\mathbf{A}}$

Act A is wrong
= $\sim R\underline{\mathbf{A}}$ = Act A isn't all right
= $O\sim\underline{\mathbf{A}}$ = Act A ought not to be done

It ought to be that A and B
= $O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

It's all right that A or B
= $R(\underline{\mathbf{A}} \vee \underline{\mathbf{B}})$

If you do A, then you ought not to do B
= $(A \supset O\sim\underline{\mathbf{B}})$

You ought not to combine doing A with doing B
= $O\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

The last pair are *deontic* if-then and don't-combine forms.

Here are translations using quantifiers:

It's obligatory that everyone do A = $O(x)A\underline{\mathbf{x}}$

[1] I'm also taking imperatives in an all-things-considered (not *prima facie*) sense. So I don't take "Do A" to mean "Other-things-being-equal, do A."

[2] For a discussion of whether moral judgments are true-or-false (as I contend they are), see my *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018) and *Ethics and Religion* (New York: Cambridge, 2016).

It's not obligatory that everyone do A = $\sim O(x)A\underline{\mathbf{x}}$
It's obligatory that not everyone do A = $O\sim(x)A\underline{\mathbf{x}}$
It's obligatory that everyone refrain from doing A = $O(x)\sim A\underline{\mathbf{x}}$

These two are importantly different:

It's obligatory that someone answer the phone = $O(\exists x)A\underline{\mathbf{x}}$

There's someone who has the obligation to answer the phone = $(\exists x)OA\underline{\mathbf{x}}$

The first might be true while the second is false; it might be obligatory (on the group) that someone or other in the office answer the phone – while yet no specific person has the obligation to answer it. To prevent the "Let the other person do it" mentality in such cases, we sometimes need to assign duties.

Compare these three:

It's obligatory that some who kill repent
= $O(\exists x)(Kx \bullet R\underline{\mathbf{x}})$

It's obligatory that some kill who repent
= $O(\exists x)(K\underline{\mathbf{x}} \bullet Rx)$

It's obligatory that some both kill and repent
= $O(\exists x)(K\underline{\mathbf{x}} \bullet R\underline{\mathbf{x}})$

These three are importantly different; underlining in the wffs shows which parts are obligatory: repenting, killing, or killing-and-repenting. If we just attached "O" to indicatives, our formulas couldn't distinguish the forms; all three would translate as "$O(\exists x)(Kx \bullet Rx)$." Because of such examples, we need to attach "O" 0278 to imperative wffs, not to indicative ones.[1]

Wffs in deontic logic divide broadly into *descriptive*, *imperative*, and *deontic* (*normative*). Here are examples of each:

- *Descriptive* ("You do A"): A, Au
- *Imperative* ("Do A"): $\underline{\mathbf{A}}$, $A\underline{\mathbf{u}}$
- *Deontic* ("ought" or "all right"): $O\underline{\mathbf{A}}$, $OA\underline{\mathbf{u}}$, $R\underline{\mathbf{A}}$, $RA\underline{\mathbf{u}}$

Such wff-types can matter for logic; for example, "O" and "R" must attach to *imperative* wffs. Here are rules for distinguishing these three types of wff:

[1] We can't distinguish the three as "$(\exists x)(Kx \bullet ORx)$," "$(\exists x)(OKx \bullet Rx)$," and "$(\exists x)O(Kx \bullet Rx)$" – since putting "$(\exists x)$" outside "O" changes the meaning. See the previous paragraph.

- Any not-underlined capital letter not immediately followed by a small letter is a descriptive wff. Any underlined capital letter not immediately followed by a small letter is an imperative wff.
- The result of writing a not-underlined capital letter and then one or more small letters, none of which are underlined, is a descriptive wff. The result of writing a not-underlined capital letter and then one or more small letters, one small letter of which is underlined, is an imperative wff.
- The result of prefixing any wff with "∼" is a wff and is descriptive, imperative, or deontic, depending on what the original wff was.
- The result of joining any two wffs by "•" or "∨" or "⊃" or "≡" and enclosing the result in parentheses is a wff. The resulting wff is descriptive if both original wffs were descriptive; it's imperative if at least one was imperative; it's deontic if both were deontic or if one was deontic and the other descriptive.
- The result of writing a quantifier and then a wff is a wff and is descriptive, imperative, or deontic, depending on what the original wff was.
- The result of writing a small letter and then "=a" and then a small letter is a descriptive wff.
- The result of writing "◇" or "☐," and then a wff, is a descriptive wff.
- The result of writing "O" or "R," and then an imperative wff, is a deontic wff.

12.3a Exercise: LogiCola L (DM & DT)

Translate these English sentences into wffs; take each "you" as a singular "you."

"You ought to do A" entails "It's possible that you do A."
☐(O**<u>A</u>** ⊃ ◇A) Here "◇A" doesn't use underlining; "◇A" means "It's possible that you do A" – while "◇A" means "The imperative 'Do A' is logically consistent." Our sentence also could translate as "☐(OA**<u>u</u>** ⊃ ◇Au)."

0279

1. If you're accelerating, then you ought not to brake. [Use A and B.]
2. You ought not to combine accelerating with braking.
3. If A is wrong, then don't do A.
4. Do A, only if A is permissible.
5. "Do A" entails "A is permissible."
6. Act A is morally indifferent (morally optional).
7. If A is permissible and B is permissible, then A-and-B is permissible.
8. It's not your duty to do A, but it's your duty not to do A.

9. If you believe that you ought to do A, then you ought to do A. [Use B for "You believe that you ought to do A" and A for "You do A."]

10. You ought not to combine believing that you ought to do A with not doing A.

11. "Everyone does A" doesn't entail "It would be all right for you to do A." [Ax, u]

12. If it's all right for X to do A to Y, then it's all right for Y to do A to X. [Axy]

13. It's your duty to do A, only if it's possible for you to do A.

14. It's obligatory that the state send only guilty persons to prison. [Gx, Sxy, s]

15. If it's not possible for everyone to do A, then you ought not to do A. [Ax, u]

16. If it's all right for someone to do A, then it's all right for everyone to do A.

17. If it's all right for you to do A, then it's all right for anyone to do A.

18. It's not all right for anyone to do A.

19. It's permissible that everyone who isn't sinful be thankful. [Sx, Tx]

20. It's permissible that everyone who isn't thankful be sinful.

12.4 Deontic proofs

We'll now add six inference rules. The first four, following the modal and quantificational pattern, are for reversing squiggles and dropping "R" and "O."

These reverse-squiggle rules hold regardless of what pair of contradictory imperative wffs replaces "$\underline{\mathbf{A}}$"/"$\sim\underline{\mathbf{A}}$":

Reverse squiggle RS
$\sim O\underline{\mathbf{A}} \rightarrow R\sim\underline{\mathbf{A}}$
$\sim R\underline{\mathbf{A}} \rightarrow O\sim\underline{\mathbf{A}}$

These let us go from "not obligatory to do" to "permissible not to do" – and from "not permissible to do" to "obligatory not to do." Use these rules only within the same world and only when the formula begins with "~O" or "~R."

We need to expand our worlds. From now on, a **possible world** is a consistent set of indicatives *and imperatives*. And a **deontic world** is a possible world (in this expanded sense) in which (a) the indicative statements are all true and (b) the imperatives prescribe some jointly permissible combination of actions. So these equivalences hold:

O**A** (A is obligatory) = ”Do A” is in *all* deontic worlds

R**A** (A is permissible) = ”Do A” is in *some* deontic worlds 0280

Suppose I have an 8 am class (C), I ought to get up before 7 am (O**G**), it would be permissible for me to get up at 6:45 am (R**A**), and it would be permissible for me to get up at 6:30 am (R**B**). Then every deontic world would have “C” and “**G**”; but some deontic worlds would have “**A**” while others would have “**B**.”

A **world prefix** now is a string of zero or more instances of “W” or “D.” As before, world prefixes represent possible worlds. “D,” “DD,” and so on represent deontic worlds; we can use these in derived lines and assumptions, such as:

D ∴ A (So A is true in deontic world D.)

DD asm: A (Assume A is true in deontic world DD.)

We can drop deontic operators using the next two rules (which hold regardless of what imperative wff replaces “**A**”). Here’s the drop-“R” rule:

Drop “R” DR
R**A** → D ∴ **A**, use a *new* string of D’s

Here the line with “R**A**” can use any world prefix – and the line with “∴ **A**” must use a world prefix that’s the same except that it ends with a *new* string (a string not occurring in earlier lines) of one or more D’s. If act A is permissible, then “Do A” is in some deontic world; we may give this world an arbitrary and hence *new* name – corresponding to a new string of D’s. We’ll use “D” for the first “R” we drop, “DD” for the second, and so forth. So if we drop two R’s, then we must introduce two deontic worlds:

R**A**
R**B**
–––––––
D ∴ **A**
DD ∴ **B**

Act A is permissible, act B is permissible; so some deontic world (call it “D”) has “Do A” and another (call it “DD”) has “Do B.” It’s OK to use “D” in the first inference, since it occurs in no earlier line; but the second inference must use

“DD,” since “D” has now already occurred. So permissible options need not be combinable; if it’s permissible to marry Ann and permissible to marry Beth, it needn’t be permissible to marry both Ann and Beth (bigamy).

We can drop an “R” from formulas that are more complicated, as long as “R” *begins* the wff; so this first inference is fine:

$R(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$
––––––––––
$D \therefore (\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

These next two examples are wrong (since the formula doesn’t begin with an “R” – instead, it begins with a left-hand parenthesis):

$(R\underline{\mathbf{A}} \supset B)$
––––––––––
$D \therefore (\underline{\mathbf{A}} \supset B)$

$(R\underline{\mathbf{A}} \bullet R\underline{\mathbf{B}})$
–––––––––
$D \therefore (\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

Drop only an *initial* “R” – and introduce a new and different deontic world whenever you drop an “R.”

Here’s the drop-“O” rule: 0281

Drop “O” DO
$O\underline{\mathbf{A}} \rightarrow D \therefore \underline{\mathbf{A}}$,
use a blank or any string of D’s

Here the line with “$O\underline{\mathbf{A}}$” can use any world prefix, and the line with “$\therefore \underline{\mathbf{A}}$” must use a world prefix which is either the same or else the same except that it adds one or more D’s at the end. If act A is obligatory, then “Do A” is in all deontic worlds. So if we have “$O\underline{\mathbf{A}}$” in the actual world, then we can derive “$\therefore \underline{\mathbf{A}}$,” “$D \therefore \underline{\mathbf{A}}$,” “$DD \therefore \underline{\mathbf{A}}$,” and so on; but it’s good strategy to stay in *old* deontic worlds when dropping “O” (and to use the actual world if there are no world with D’s). As before, we can drop an “O” from formulas that are more complicated, as long as “O” *begins* the wff. So this next inference is fine:

$O(A \supset \underline{\mathbf{B}})$
––––––––––
$D \therefore (A \supset \underline{\mathbf{B}})$

These next two example are wrong (since the formula doesn't begin with "O" – instead it begins with a left-hand parenthesis – *drop only initial operators*):

$(O\underline{\mathbf{A}} \supset B)$
––––––––––
$D \therefore (\underline{\mathbf{A}} \supset B)$

$(O\underline{\mathbf{A}} \supset O\underline{\mathbf{B}})$
––––––––––
$D \therefore (\underline{\mathbf{A}} \supset \underline{\mathbf{B}})$

"$(O\underline{\mathbf{A}} \supset B)$" and "$(O\underline{\mathbf{A}} \supset O\underline{\mathbf{B}})$" are if-then forms and follow the if-then rules: if we have the first part true, we can get the second true; if we have the second part false, we can get the first false; and if we get stuck, we'll need to make another assumption.

Rule DO lets us go from "$O\underline{\mathbf{A}}$" in a world to "$\underline{\mathbf{A}}$" in the same world. This accords with "Hare's Law" (named after R. M. Hare):

Hare's Law

$\Box(O\underline{\mathbf{A}} \supset \underline{\mathbf{A}})$

An ought judgment entails the corresponding imperative: "You ought to do A" entails "Do A."

Hare's Law (also called "prescriptivity") equivalently claims that "You ought to do it, but don't" is inconsistent. This law fails for some weaker *prima facie* or descriptive senses of "ought"; there's no inconsistency in this: "You ought (according to company policy) to do it, but don't do it." The law seems to hold for the all-things-considered, normative sense of "ought"; this seems inconsistent: "All things considered, you ought to do it; but don't do it." However, some philosophers reject Hare's Law; those who reject it would want to specify that in applying rule DO the world prefix of the derived line has to end in a "D" (and so we can't use a blank world prefix in the derived line).

Here's a deontic proof using these rules: 0282

1 $O\sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ **Valid**
2 $O\underline{\mathbf{A}}$
[$\therefore O\sim\underline{\mathbf{B}}$
* 3 ┌ asm: $\sim O\sim\underline{\mathbf{B}}$
* 4 │ $\therefore R\underline{\mathbf{B}}$ {from 3}
5 │ $D \therefore \underline{\mathbf{B}}$ {from 4}
* 6 │ $D \therefore \sim(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$ {from 1}
7 │ $D \therefore \underline{\mathbf{A}}$ {from 2}

8 └ D ∴ ~<u>**B**</u>   {from 6 and 7}
9 ∴ O~<u>**B**</u>   {from 3; 5 contradicts 8}

Reverse a squiggle (line 4). Drop an initial "R," using a new deontic world (line 5). Drop each initial "O," using the same old deontic world (lines 6 and 7). This all works like a modal proof, except for underlining and having "O," "R," and "D" in place of "☐," "◇," and "W." As with modal logic, we can star (and then ignore) a line when we use a reverse-squiggle or "R"-dropping rule on it.

Things get more complicated if we use the rules for dropping "R" and "O" on a formula in some other possible world. Here's a simple case. Formulas "R<u>**A**</u>" and "O<u>**B**</u>" are in the actual world (using the blank world prefix); and so we put the corresponding imperatives in a deontic world "D."

R<u>**A**</u>
O<u>**B**</u>

D ∴ <u>**A**</u>
D ∴ <u>**B**</u>

In the next example, formulas "R<u>**A**</u>" and "O<u>**B**</u>" are in world W; so here we keep "W" and just add "D" (the rules for dropping "R" and "O" allow these moves):

W ∴ R<u>**A**</u>
W ∴ O<u>**B**</u>

WD ∴ <u>**A**</u>
WD ∴ <u>**B**</u>

Here world WD is a deontic world that *depends on* possible world W; this means that (a) the indicative statements in WD are those of world W, and (b) the imperatives of WD prescribe some set of actions that are jointly permissible according to the deontic judgments of world W. The following proof uses world prefix "WD" in lines 7 to 9:

[∴ ☐(O(<u>**A**</u> • <u>**B**</u>) ⊃ O<u>**A**</u>)   **Valid**
* 1 ┌ asm: ~☐(O(<u>**A**</u> • <u>**B**</u>) ⊃ O<u>**A**</u>)
* 2 │ ∴ ◇~(O(<u>**A**</u> • <u>**B**</u>) ⊃ O<u>**A**</u>)   {from 1}
* 3 │ W ∴ ~(O(<u>**A**</u> • <u>**B**</u>) ⊃ O<u>**A**</u>)   {from 2}
4 │ W ∴ O(<u>**A**</u> • <u>**B**</u>)   {from 3}
* 5 │ W ∴ ~O<u>**A**</u>   {from 3}
* 6 │ W ∴ R~<u>**A**</u>   {from 5}
7 │ WD ∴ ~<u>**A**</u>   {from 6}
8 │ WD ∴ (<u>**A**</u> • <u>**B**</u>)   {from 4}

9 └ WD ∴ **A** {from 8}
10 ∴ ☐(O(**A** • **B**) ⊃ O**A**) {from 1; 7 contradicts 9}

0283 When we drop the "R" from line 6 ("W ∴ R∼**A**"), we add a new deontic world D to world W, so we get "WD ∴ ∼**A**" in line 7. The next two chapters will often use complex world prefixes like "WD."

We have two more rules. The indicative-transfer rule lets us transfer indicatives freely between a deontic world and whatever world it depends on; we can do this because these two worlds have the same indicative (descriptive or deontic) wffs. IT holds regardless of what descriptive or deontic wff replaces "A":

Indicative transfer IT
D ∴ A → A

The world prefixes of the derived and deriving lines must be identical except that one ends in one or more additional D's. Here are some correct uses:

A
——————
D ∴ A

D ∴ A
———————
∴ A

D ∴ A
———————
DD ∴ A

O**A**
———————
D ∴ O**A**

This next inference is wong, since IT is to be used only with indicatives (including deontic judgments):

A
—————
D ∴ **A**

It can be useful to move an indicative between deontic worlds when we need to do so to get a contradiction or apply an I-rule. Here's an example:

It’s obligatory that all teachers prepare classes.
You’re a teacher.
∴ You ought to prepare classes.

1 O(x)(Tx ⊃ P$\underline{x}$) **Valid**
2 Tu
[∴ OP$\underline{u}$
* 3 ┌ asm: ~OP$\underline{u}$
* 4 │ ∴ R~P$\underline{u}$ {from 3}
5 │ D ∴ ~P$\underline{u}$ {from 4}
* 6 │ D ∴ (x)(Tx ⊃ P$\underline{x}$) {from 1}
7 │ D ∴ (Tu ⊃ P$\underline{u}$) {from 6}
8 │ D ∴ ~Tu {from 5 and 7}
9 └ D ∴ Tu {from 2}
10 ∴ OP$\underline{u}$ {from 3; 8 contradicts 9}

Instead of moving the indicative “Tu” from the actual world to D in line 9, we could have moved “~Tu” from D to the actual world.

Our final inference rule, **Kant’s Law**, is named for Immanuel Kant:

Kant’s Law KL
O$\underline{A}$ → ◇A
“Ought” implies “can”: “You ought to do A” entails “It’s possible for you to do A.”

This holds regardless of what imperative wff replaces “**$\underline{A}$**” and what indicative wff replaces “A,” if the former is like the latter except for underlining, and every 0284 wff out of which the former is constructed is an imperative.[1] Kant’s Law is often useful with arguments having both deontic (“O” or “R”) and modal operators (“☐” or “◇”); note that you infer “◇A” (“It’s possible for you to do A”) and not “◇**$\underline{A}$**” (“The imperative ‘Do A’ is consistent”).

Kant’s Law equivalently claims that “You ought to do it, but it’s impossible” is inconsistent. This law fails for some weaker *prima facie* or descriptive senses of “ought”; since company policy may require impossible things, this is consistent: “You ought (according to company policy) to do it, but it’s impossible.” The law seems to hold for the all-things-considered, normative sense of “ought”; this seems inconsistent: “All things considered, you ought to do it; but it’s impossible to do it.” We can’t have an all-things-considered moral

[1] The proviso outlaws “O(∃x)(Lx • ~L$\underline{x}$) ∴ ◇(∃x)(Lx • ~Lx)” (“It’s obligatory that someone who is lying not lie ∴ It’s possible that someone both lie and not lie”). Since “L**x**” in the premise isn’t an imperative wff, this (incorrect) derivation doesn’t satisfy KL.

obligation to do the impossible.

KL is a weak form of Kant's Law. Kant thought that what we ought to do is not just *logically possible*, but also what we're *capable of doing* (physically and psychologically). Our rule KL expresses only the "logically possible" part; but, even so, it's still useful for many arguments. And it won't hurt if sometimes we informally interpret "$\Diamond$" in terms of what we're *capable of doing*.

We've already mentioned the first two of these four "laws":[1]

- **Hare's Law**: An "ought" entails the corresponding imperative.
- **Kant's Law**: "Ought" implies "can."
- **Hume's Law**: We can't deduce an "ought" from an "is."
- **Poincaré's Law**: We can't deduce an imperative from an "is."

Now we'll briefly consider the last two.

Hume's Law (named for David Hume) claims that we can't validly deduce what we *ought* to do from premises that don't contain "ought" or similar notions.[2] So getting a moral conclusion requires having a moral premise. Hume's Law fails for some weak senses of "ought"; given descriptions of company policy and the situation, we can sometimes validly deduce what ought (according to company policy) to be done. Hume's Law seems to hold for the all-things-considered, normative sense of "ought." A more careful wording would say: "If B is a consistent non-evaluative statement and A a simple contingent action, then B doesn't entail 'Act A ought to be done.'" This wording sidesteps some counterexamples (§12.4a) where we clearly *can* deduce an "ought" from an "is."

Poincaré's Law (named for the mathematician Jules Henri Poincaré) similarly claims that we can't validly deduce an imperative from indicative premises that 0285 don't contain "ought" or similar notions. A more careful wording would say: "If B is a consistent non-evaluative statement and A a simple contingent action, then B doesn't entail the imperative 'Do act A.'" Again, the qualifications block objections (like problems 9 and 10 of §12.2b). We won't build Hume's or Poincaré's Law into our system.

Our deontic proof strategy is much like the modal strategy. First we reverse squiggles to put "O" and "R" at the beginning of a formula. Then we drop each initial "R," putting each permissible thing into a *new* deontic world. Lastly we drop each initial "O," putting each obligatory thing into each *old* deontic world. Drop obligatory things into the actual world just if:

[1] The word "law," although traditional here, is really too strong, since all four are controversial and subject to qualifications.

[2] Some philosophers disagree and claim we can deduce moral conclusions using only premises about social conventions, personal feelings, God's will, or something similar. For views on both sides, see my *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018).

- the premises or conclusion have an instance of an underlined letter that isn't part of some wff beginning with "O" or "R"; or
- you've done everything else possible (including further assumptions if needed) and still have no old deontic worlds.

Use the indicative transfer rule if you need to move an indicative between the actual world and a deontic world (or vice versa). Consider using Kant's Law if you see a letter that occurs underlined in a deontic wff and not-underlined in a modal wff; some proofs that use Kant's Law get tricky.

From now on, we won't do refutations for invalid arguments in the book (LogiCola keeps doing them), since refutations get too messy when we mix various kinds of world.

12.4a Exercise: LogiCola M (D & M)

Say whether valid (and give a proof) or invalid (no refutation necessary).

∴ ~◇(O**A** • O~**A**)

[∴ ~◇(O**A** • O~**A**)   **Valid**

* 1 ┌ asm: ◇(O**A** • O~**A**)

* 2 │ W ∴ (O**A** • O~**A**)   {from 1}

3 │ W ∴ O**A**   {from 2}

4 │ W ∴ O~**A**   {from 2}

5 │ W ∴ **A**   {from 3}

6 └ W ∴ ~**A**   {from 4}

7 ∴ ~◇(O**A** • O~**A**)   {from 1; 5 contradicts 6}

This wff says "It's not logically possible that you ought to do A and also ought not to do A"; this is correct if we take "ought" in the all-things-considered, normative sense. Morality can't make impossible demands on us; if we think otherwise, our lives will likely be filled with irrational guilt for not fulfilling impossible demands. "~◇(O**A** • O~**A**)" would be incorrect if we took "O" in it to mean something like "ought according to company policy" or "prima facie ought." Inconsistent company policies may require that we do A and also require that we not do A; and we can have a prima facie duty to do A and another to omit doing A.

0286

1. O~**A**
∴ O~(**A** • **B**)

2. (∃x)OA**x**
∴ O(∃x)A**x**

3. $b=c$
$\therefore (OF\underline{\mathbf{a}}b \supset OF\underline{\mathbf{a}}c)$

4. $\therefore O(O\underline{\mathbf{A}} \supset \underline{\mathbf{A}})$

5. $\therefore O(\underline{\mathbf{A}} \supset O\underline{\mathbf{A}})$

6. $\therefore O(\underline{\mathbf{A}} \supset R\underline{\mathbf{A}})$

7. $O\underline{\mathbf{A}}$
$O\underline{\mathbf{B}}$
$\therefore O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

8. $(x)OF\underline{\mathbf{x}}$
$\therefore O(x)F\underline{\mathbf{x}}$

9. $O(\underline{\mathbf{A}} \vee \underline{\mathbf{B}})$
$\therefore (\sim\Diamond A \supset R\underline{\mathbf{B}})$

10. $(A \supset O\underline{\mathbf{B}})$
$\therefore O(A \supset \underline{\mathbf{B}})$

11. $\Box(\underline{\mathbf{A}} \supset \underline{\mathbf{B}})$
$O\underline{\mathbf{A}}$
$\therefore O\underline{\mathbf{B}}$

12. $O\underline{\mathbf{A}}$
$R\underline{\mathbf{B}}$
$\therefore R(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$

13. A
$\therefore O(\underline{\mathbf{B}} \vee \sim\underline{\mathbf{B}})$

14. $(x)RA\underline{\mathbf{x}}$
$\therefore R(x)A\underline{\mathbf{x}}$

15. $O\underline{\mathbf{A}}$
$O\underline{\mathbf{B}}$
$\therefore \Diamond(A \bullet B)$

16. $\therefore (R\underline{\mathbf{A}} \vee R\sim\underline{\mathbf{A}})$

17. (O**A** ⊃ **B**)
∴ R(**A** • **B**)

18. ~◇A
∴ R~**A**

19. A
~A
∴ O**B**

20. O(x)(Fx ⊃ G**x**)
OF**a**
∴ OG**a**

21. O(A ⊃ **B**)
∴ (A ⊃ O**B**)

22. O(x)A**x**
∴ (x)OA**x**

23. ∴ O(~R**A** ⊃ ~**A**)

24. A
∴ (A ∨ O**B**)

25. (A ∨ O**B**)
~A
∴ O**B**

Problems 3, 13, and 19 deduce an "ought" from an "is." If "(A ∨ O**B**)" is an "ought," then 24 is another example; if it's an "is," then 25 is another example. 20 of §12.4b is another example. We formulated Hume's Law so that these examples don't refute it.

12.4b Exercise: LogiCola M (D & M)

First appraise intuitively. Then translate into logic (using the letters given) and say whether valid (and give a proof) or invalid (no refutation necessary).

1. It's not all right for you to combine texting with driving.
You ought to drive.
∴ Don't text. [Use T and D.]

2. ∴ Either it's your duty to do A or it's your duty not to do A. [The conclusion, if taken to apply to every action A, is *rigorism*, the view that there are no morally neutral acts (acts permissible to do and also permissible not to do).]

3. I did A.
I ought not to have done A.
If I did A and it was possible for me not to have done A, then I have free will.
∴ I have free will. [Use A and F. Immanuel Kant thus argued that ethics requires free will.]

4. ∴ If you ought to do A, then do A.

5. ∴ If you ought to do A, then you'll in fact do A. 0287

6. It's not possible for you to be perfect.
∴ It's not your duty to be perfect. [Use "P" for "You are perfect."]

7. You ought not to combine texting with driving.
You don't have a duty to drive.
∴ It's all right for you to text. [T, D]

8. ∴ Do A, only if it would be all right for you to do A.

9. If it's all right for you to insult Jones, then it's all right for Jones to insult you.
∴ If Jones ought not to insult you, then don't you insult Jones. [Use Ixy, u, and j. The premise follows from the universalizability principle ("What's right for one person is right for anyone else in similar circumstances") plus the claim that the cases are similar. The conclusion is a distant relative of the golden rule.]

10. It's all right for someone to do A.
∴ It's all right for anyone to do A. [Can you think of an example where the premise would be true and conclusion false?]

11. If fatalism (the view that whatever happens couldn't have been otherwise) is true and I do A, then my doing A (taken by itself) is necessary.
∴ If fatalism is true and I do A, then it's all right for me to do A. [F, A]

12. If it's all right for you to complain, then you ought to take action.
∴ You ought to either take action or else not complain. [Use C and T. This is the "Put up or shut up" argument.]

13. I ought to stay with my brother while he's sick in bed.
It's impossible for me to combine these two things: staying with my brother while he's sick in bed and driving you to the airport.
∴ It's all right for me not to drive you to the airport.   [S, D]

14. Jones ought to be happy in proportion to his moral virtue.
Necessarily, if Jones is happy in proportion to his moral virtue, then Jones will be rewarded either in the present life or in an afterlife.
It's not possible for Jones to be rewarded in the present life.
If it's possible for Jones to be rewarded in an afterlife, then there is a God.
∴ There is a God.   [Use H for "Jones is happy in proportion to his moral virtue," P for "Jones will be rewarded in the present life," A for "Jones will be rewarded in an afterlife," and G for "There is a God." This is Kant's moral argument for the existence of God. To make premise 3 plausible, we must take "possible" as "factually possible" (instead of "logically possible"). But does "ought to be" (premise 1 uses this – and not "ought to do") entail "is factually possible"?]

15. If killing the innocent is wrong, then one ought not to intend to kill the innocent.
If it's permissible to have a nuclear retaliation policy, then intending to kill the innocent is permissible.
∴ If killing the innocent is wrong, then it's wrong to have a nuclear retaliation policy.   [K, I, N] 0288

16. If it's all right for you to do A, then you ought to do A.
If you ought to do A, then it's obligatory that everyone do A.
∴ If it's impossible that everyone do A, then you ought not to do A.   [Use Ax and u. The premises and conclusion are doubtful; the conclusion entails "If it's impossible that everyone become the first woman president, then you ought not to become the first woman president." The conclusion is a relative of Kant's formula of universal law; it's also a faulty "formal ethical principle" – an ethical principle that we can formulate using abstract logical notions but leaving unspecified the meaning of the individual, property, relational, and statement letters.]

17. It's obligatory that Smith help someone or other whom Jones is beating up.
∴ It's obligatory that Jones beat up someone.   [Use Hxy, Bxy, s, and j. This "good Samaritan paradox" is provable in most deontic systems that attach "O" to indicatives. There are similar examples where the evil deed happens after the good one. It may be obligatory that Smith warn someone or other whom Jones will try to beat up; this doesn't entail that Jones ought to try to beat up someone.]

18. If it's not right to do A, then it's not right to promise to do A.
∴ Promise to do A, only if it's all right to do A.   [A, P]

19. It's obligatory that someone answer the phone.
∴ There's someone who has the obligation to answer the phone. [Ax]

20. Studying is needed to become a teacher.
"Become a teacher" entails "Do what's needed to become a teacher."
"Do what's needed to become a teacher" entails "If studying is needed to become a teacher, then study."
∴ You ought to either study or not become a teacher. [Use N for "Studying is needed to become a teacher," B for "You become a teacher," D for "You do what's needed to become a teacher," and S for "You study." This is an ought-version of §12.2b #9. It shows that we can deduce a complex ought judgment from purely descriptive premises.]

21. If it's right for you to litter, then it's wrong for you to preach concern for the environment.
∴ It's not right for you to combine preaching concern for the environment with littering. [L, P]

22. If you ought to be better than everyone else, then it's obligatory that everyone be better than everyone else.
"Everyone is better than everyone else" is self-contradictory.
∴ It's all right for you not to be better than everyone else. [Use Bx (for "x is better than everyone else") and u.]

23. You ought not to combine braking with accelerating.
You ought to brake.
∴ You ought to brake and not accelerate. [B, A] 0289

24. "Everyone breaks promises" is impossible.
∴ It's all right for there to be someone who doesn't break promises. [Use Bx. Kant thought universal promise-breaking would be impossible, since no one would make promises if everyone broke them. But he wanted to draw the stronger conclusion that it's always wrong to break promises. See problem 16.]

25. It's all right for you to punish Judy for the accident, only if Judy ought to have stopped her car more quickly.
Judy couldn't have stopped her car more quickly.
∴ You ought not to punish Judy for the accident. [P, S]

26. You ought to pay by check or pay by MasterCard.
If your MasterCard is expired, then you ought not to pay by MasterCard.
∴ If your MasterCard is expired, then pay by check. [C, M, E]

27. You ought to help your neighbor.
It ought to be that, if you (in fact) help your neighbor, then you say you'll help him.
You don't help your neighbor.
If you don't help your neighbor, then you ought not to say you'll help him.
∴ You ought to say you'll help him, and you ought not to say you'll help him. [Use H and S. Roderick Chisholm pointed out that this clearly invalid argument was provable in many systems of deontic logic. Is it provable in our system?]

28. If you take logic, then you'll make mistakes.
You ought not to make mistakes.
∴ You ought not to take logic.   [T, M]

29. If I ought to name you acting mayor because you served on the city council, then I ought to name Jennifer acting mayor because she served on the city council.
I can't name both you and Jennifer acting mayor.
∴ It's false that I ought to name you acting mayor because you served on the city council.   [U, J]

13 Belief Logic

Our **belief logic** is "logic" in an extended sense. Instead of studying what follows from what, it studies patterns of consistent believing and willing; it generates consistency norms that prescribe that we be consistent in various ways. We'll start with a simplified system and then add refinements.

13.1 Belief translations

We'll use ":" to construct descriptive and imperative belief formulas:

1. The result of writing a small letter and then ":" and then a wff is a descriptive wff.
2. The result of writing an underlined small letter and then ":" and then a wff is an imperative wff.

Statements about beliefs translate into *descriptive* belief formulas:

You believe that A is true = u:A

You don't believe that A is true = ∼u:A

You believe that A is false = u:∼A

You don't believe A and you don't believe not-A
= (∼u:A • ∼u:∼A)

If you refrain from believing A, you might believe that A is false or you might take no position on A. Here are some further translations:

You believe that you ought to do A
= u:OA$\underline{\mathbf{u}}$

Everyone believes that they ought to do A
= (x)x:OA$\underline{\mathbf{x}}$

You believe that if A then not-B
= u:(A ⊃ ∼B)

If you believe A, then you don't believe B
= $(u{:}A \supset \sim u{:}B)$

Since our belief logic generates norms *prescribing* consistency, it focuses on *imperative* belief formulas – which we express by underlining the small letter: 0291

Believe that A is true = $\underline{\mathbf{u}}{:}A$

Don't believe that A is true = $\sim\underline{\mathbf{u}}{:}A$

Believe that A is false = $\underline{\mathbf{u}}{:}\sim A$

Don't believe A and don't believe not-A
= $(\sim\underline{\mathbf{u}}{:}A \bullet \sim\underline{\mathbf{u}}{:}\sim A)$

Believe that you ought to do A
= $\underline{\mathbf{u}}{:}OA\underline{\mathbf{u}}$

Let everyone believe that they ought to do A
= $(x)\underline{\mathbf{x}}{:}OA\underline{\mathbf{x}}$

As before, we distinguish between if-then and don't-combine forms:

If you in fact believe A, then don't believe B
= $(u{:}A \supset \sim\underline{\mathbf{u}}{:}B)$

Don't combine believing A with believing B
= $\sim(\underline{\mathbf{u}}{:}A \bullet \underline{\mathbf{u}}{:}B)$

13.1a Exercise: LogiCola N (BM & BT)

Translate these sentences into wffs (use "u" for "you" and "G" for "There's a God").

You believe that there's a God. (You're a theist.)
u:G

1. You believe that there's no God. (You're an atheist.)
2. You take no position on whether there's a God. (You're an agnostic.)
3. You don't believe that there's a God. (You're a non-theist.)
4. You believe that "There's a God" is self-contradictory.
5. Necessarily, if you're a theist then you aren't an atheist. (Is this statement true?)

6. Believe that there's a God.
7. If "There's a God" is self-contradictory, then don't believe that there's a God.
8. If you believe A, then you don't believe not-A.
9. If you believe A, then don't believe not-A.
10. Don't combine believing A with believing not-A.

13.2 Belief proofs

There are three approaches to belief logic. First, we might study what belief formulas validly follow from what other belief formulas. We might try to prove arguments like this one:

You believe A.
∴ You don't believe not-A.

u:A
∴ ~u:~A

But this is invalid, since people can be confused and illogical. Students and politicians can assert A and assert not-A almost in the same breath. Beginning ethics students often write things like this (§4.3): 0292

Since morality is relative to culture, no duties bind universally. What's right in one culture is wrong in another. Universal duties are a myth. Relativism should make us tolerant toward others; we can't say that we're right and they're wrong. So everyone ought to respect the values of others.

Here "No duties bind universally" clashes with "Everyone ought to respect the values of others." As Socrates was adept at showing, our unexamined views are often filled with inconsistencies. But then, given that someone believes A, we can *deduce* little or nothing about what else the person believes or doesn't believe. So this first approach to belief logic is doomed to failure.

A second approach studies *how we'd believe if we were completely consistent*. A person X is **completely consistent** (an idealized notion) if and only if:

1. the set S of things that X believes is logically consistent, and
2. X believes whatever follows logically from set S.

Our previous argument would be valid if we added, as an additional premise, that you're completely consistent:

You're completely consistent. (implicit)
You believe A.
∴ You don't believe not-A.

Belief logic would take "You're completely consistent" as an implicit premise; this would be assumed, even though it's false, to help us explore what belief patterns a consistent person would follow. While this works,[1] I prefer a third approach, in view what I want to do in the next chapter.

My third approach generates consistency imperatives, like these:

Don't combine believing A with believing not-A.
$\sim(\underline{\mathbf{u}}:A \bullet \underline{\mathbf{u}}:\sim A)$

Don't combine believing A-and-B with not believing A.
$\sim(\underline{\mathbf{u}}:(A \bullet B) \bullet \sim\underline{\mathbf{u}}:A)$

This third approach will assume that *we ought to be consistent* – we ought not to combine inconsistent beliefs and we ought not to believe something without also believing whatever follows from it. While this basic idea is plausible (but subject to qualifications, see §13.7), it's not easy to systematize logically.

Our belief logic adds belief worlds and inference rules to our proof machinery. We represent a *belief world* by a string of one or more instances of a small-letter constant. Since most of our belief norms use a generic "you," our belief worlds will typically be "u," "uu," "uuu," and so on. So a **world prefix** is now a string of zero or more instances of letters from the set <W, D, a, b, c, …>, where <a, b, c, …> is the set of small-letter constants. Our two inference rules use belief 0293 worlds; while it's fairly easy to use these rules mechanically, it's difficult to get an intuitive grasp of how they work. Let me try to explain them.

First, let a *belief policy* be a set of imperatives about what someone (typically a generic "you") is or is not to believe. Here's an example:

[1] Jaakko Hintikka used roughly this second approach in his classic *Knowledge and Belief* (Ithaca, New York: Cornell University Press, 1962).

Believe that Michigan will play.

$\underline{\mathbf{u}}$:P

Be neutral about whether Michigan will win.

($\sim\underline{\mathbf{u}}$:W • $\sim\underline{\mathbf{u}}$:~W)

This policy prescribes a way to believe that's consistent (but boring). In general, a belief policy prescribes a *consistent way to believe* if and only if (1) the set S of things that the person is told to believe is logically consistent, and (2) the person isn't forbidden to believe something that follows logically from set S. Our task here is to express this idea using possible worlds. I want to reject belief policies, such as this one, that prescribe an *inconsistent way to believe*:

Believe A and believe not-A.

($\underline{\mathbf{u}}$:A • $\underline{\mathbf{u}}$:~A)

How do we reject such policies using possible worlds?

A **belief world** (relative to a belief policy about what a person is told to believe) is a possible world that contains all the statements that the person is told to believe. So if you're told to believe A, then all your belief worlds have A. Individual belief worlds may contain further statements. For example, if you're told to be neutral about B (not to believe B and not to believe not-B), then some of your belief worlds will have B and some will have not-B. What's common to all your belief worlds is what you're told to believe. If a belief policy (about what you're told to believe) forces a belief world to be self-contradictory, then the belief policy tells you to believe inconsistently; and then (by an implicit "Be consistent" built into the system) we reject the belief policy.

Our first inference rule, B+, says that, if you're told to believe A, then A is in all your belief worlds: u, uu, uuu, and so on. Rule B+ operates on *positive imperative belief formulas*; here any wff can replace "A" and any small letter can replace "u":

B+

$\underline{\mathbf{u}}$:A → u ∴ A,
use any string of u's

The line with “$\underline{\mathbf{u}}$:A” can use any world prefix with no small letters or “W”[1] – 0294 and the line with “u ∴ A” must use a world prefix that’s the same except that it adds at the end a string of one or more instances of “u” (or of the small letter that replaces “u”). If we have “u ∴ A” in a proof, “u” refers to a belief world based on what you’re told to believe. (If instead we have “Du ∴ A,” then we have a belief world based on what you’re told to believe in deontic world D.)

We can use B+ to prove this consistency imperative: “Don’t combine believing A with believing not-A.” First assume its opposite: “Believe A and believe not-A.” Then use B+ to construct a belief world that contains everything that you’re told to believe. Since this world necessarily has contradictions, “Believe A and believe not-A” tells us to believe inconsistently; then (by an implicit “Be consistent” built into the system) we can derive the opposite: “Don’t combine believing A and believing not-A.” Here’s the proof in symbols:

```
   [ ∴ ~(u:A • u:~A)    Valid
 * 1 ┌ asm: (u:A • u:~A)
   2 │ ∴ u:A    {from 1}
   3 │ ∴ u:~A    {from 1}
   4 │ u ∴ A    {from 2}
   5 └ u ∴ ~A    {from 3}
   6 ∴ ~(u:A • u:~A)    {from 1; 4 contradicts 5}
```

B+ puts the statements you’re told to believe into belief world u. Since world u has contradictions, our assumption prescribes an inconsistent combination of belief attitudes. So we reject it and derive the original conclusion.[2]

We defined “X is completely consistent” using two clauses:

1. the set S of things that X believes is logically consistent, and
2. X believes whatever follows logically from set S.

While B+ captures the first clause, we need rule B– to capture the second. By B–, if you’re told NOT to believe A, then not-A must be in SOME of your belief worlds. So if you’re told to be neutral about A (NOT to believe A and NOT to believe not-A) then some of your belief worlds will have A and some will have not-A. Rule B– operates on *negative imperative belief formulas*; any pair of contradictory wffs can replace “A” / “~A” and any small letter can replace “u”:

[1] This proviso (about small letters and “W”) blocks proofs of questionable wffs that place one imperative belief operator within another, like “$\underline{b}$:~($\underline{c}$:A • $\underline{c}$:~A),” or claim logical necessity for consistency imperatives, like “☐~($\underline{u}$:A • $\underline{u}$:~A).”

[2] Our proof doesn’t show that this conclusion is logically necessary; instead, it shows that it follows from an implicit “One ought to be consistent” premise.

B–
∼**u**:A → u ∴ ∼A, use a *new* string of u's

The line with "∼**u**:A" can use any world prefix not containing small letters or "W" – and the line with "u ∴ ∼A" must use a world prefix that's the same except that it ends with a *new* string (one not occurring in earlier lines) of one or more 0295 instances of "u" (or of the small letter that replaces "u").

We need B– to prove this consistency imperative: "Don't combine believing A-and-B with not believing A. First assume its opposite: "Believe A-and-B, but don't believe A" – which tells us to believe something but not what logically follows from it. Here's the proof:

[∴ ∼(**u**:(A • B) • ∼**u**:A)   **Valid**
* 1 asm: (**u**:(A • B) • ∼**u**:A)
2 ∴ **u**:(A • B)   {from 1}
* 3 ∴ ∼**u**:A   {from 1}
4 u ∴ ∼A   {from 3}
5 u ∴ (A • B)   {from 2}
6 u ∴ A   {from 5}
7 ∴ ∼(**u**:(A • B) • ∼**u**:A)   {from 1; 4 contradicts 6}

By B–, since you're told NOT to believe A, we put "∼A" into new belief world u (line 4). We put what you're positively told to believe into the same belief world u and then get a contradiction. Our assumption prescribes an inconsistent combination of belief attitudes. So we derive the original conclusion.

Our proof strategy goes as follows:

- First use rule B– on *negative imperative belief formulas* (formulas that say to *refrain* from believing something). Use a *new belief world* each time. You can star (and then ignore) a line when you use B– on it.
- Then use B+ on *positive imperative belief formulas* (formulas that say to believe something). Use *each old belief world* of the person in question each time. (Use a single new belief world if you have no old ones.) Don't star a line when you use B+ on it.

Both rules operate only on *imperative* belief formulas (like "∼**u**:A" or "**u**:A") – not on *descriptive* ones (like "∼u:A" or "u:A"). Our belief worlds are about what a belief policy *tells* you to believe, not about what you *actually* believe. Our proof structure is designed to prove consistency norms.

Our recent systems had rules for reversing squiggles; for dropping weak operators (*some*, *possible*, *permissible*); and for dropping strong operators (*all*, *necessary*, *ought*). Belief logic is different, since there's no convenient weak operator to go with "You believe that A" (the weak operator would have to mean "You don't believe that not-A"). Belief logic is like a modal logic with "☐" but no "◇": besides having the drop-box rule for "☐A," we'd then need a rule saying that from "∼☐A" we can put "∼A" into a new world W (like B–).

Our consistency norms have a don't-combine form, forbidding inconsistent combinations. They tell you to make your beliefs coherent with each other; but they don't say what beliefs to add or subtract to bring this about. Suppose that P (*premise*) logically entails C (*conclusion*); compare these three forms:
0296

- $(u{:}P \supset \underline{u}{:}C)$ If you believe *premise*, then believe *conclusion*
- $({\sim}u{:}C \supset {\sim}\underline{u}{:}P)$ If you don't believe *conclusion*, then don't believe *premise*
- ${\sim}(\underline{u}{:}P \bullet {\sim}\underline{u}{:}C)$ Don't combine believing *premise* with not believing *conclusion*

Suppose you believe *premise* but don't believe *conclusion*; then you violate all three. What should you do? The first form tells you to believe *conclusion*; but maybe *conclusion* is irrational and you should reject both premise and conclusion. The second tells you to drop *premise*; but maybe *premise* is solid and you should accept both premise and conclusion. So the first two forms can guide you wrongly. The third is better; it simply forbids the inconsistent combination of believing *premise* but not believing *conclusion* – but it doesn't say what to do if you get into this forbidden combination.

Here's another example. Assume that A is logically inconsistent with B; compare these three forms:

- $(u{:}A \supset {\sim}\underline{u}{:}B)$ If you believe A, then don't believe B.
- $(u{:}B \supset {\sim}\underline{u}{:}A)$ If you believe B, then don't believe A.
- ${\sim}(\underline{u}{:}A \bullet \underline{u}{:}B)$ Don't combine believing A with believing B.

Suppose you believe A and also believe B, even though the two are inconsistent. The first form tells you to drop B, while the second tells you to drop A; but which you should drop depends on the situation. The last form is better; it simply tells you to avoid the inconsistent combination.

Proofs with multiple kinds of operator can be confusing. This chart tells what order to use in dropping operators:

First drop these weak operators:

◇ ~**u**: R (∃x)

Use new worlds/constants; star the old line.

Then drop these strong operators:

☐ **u**: O (x)

Use old worlds/constants if you have them; don't star the old line.

Within each group, the dropping order doesn't matter – except that it's wise to drop "**u**:" and "O" before dropping the very strong "☐."

Section 9.2 noted that our substitute-equals rule can fail in arguments about beliefs. Consider this argument:

Jones believes that Lincoln is on the penny.
Lincoln is the first Republican president.
∴ Jones believes that the first Republican president is on the penny.

j:Pl
l=r
∴ j:Pr

If Jones is unaware that Lincoln was the first Republican president, the premises could be true while the conclusion is false. So the argument is invalid. But yet 0297 we can derive the conclusion from the premises using our substitute-equals rule. So we need to qualify this rule so it doesn't apply in belief contexts. From now on, the substitute-equals rule holds only if no interchanged instance of the constants occurs within a wff that begins with a small letter (underlined or not) followed by a colon (":").

13.2a Exercise: LogiCola OB

Say whether valid (and give a proof) or invalid (no refutation necessary).

☐(A ⊃ B)
∴ (u:A ⊃ **u**:B)

1 $\Box(A \supset B)$ **Invalid**
[$\therefore (u{:}A \supset \underline{u}{:}B)$
* 2 asm: $\sim(u{:}A \supset \underline{u}{:}B)$
3 $\therefore u{:}A$ {from 2}
* 4 $\therefore \sim\underline{u}{:}B$ {from 2}
5 $u \therefore \sim B$ {from 4}
* 6 $u \therefore (A \supset B)$ {from 1}
7 $u \therefore \sim A$ {from 5 and 6}

Since rules B+ and B– work only on *imperative* belief formulas, we can't go from "u:A" in line 3 to "u ∴ A." The conclusion here has the faulty if-then form. Suppose that A entails B and you believe A; it doesn't follow that you should believe B – maybe you should reject A and also reject B.

1. $\sim\Diamond(A \bullet B)$
$\therefore \sim(\underline{u}{:}A \bullet \underline{u}{:}B)$

2. $\sim\Diamond(A \bullet B)$
$\therefore (u{:}A \supset \sim u{:}B)$

3. $\sim\Diamond(A \bullet B)$
$\therefore (u{:}A \supset \sim\underline{u}{:}B)$

4. $\sim\Diamond(A \bullet B)$
$\therefore (\sim\underline{u}{:}A \vee \sim\underline{u}{:}B)$

5. $\sim\Diamond(A \bullet B)$
$\therefore (\underline{u}{:}\sim A \vee \underline{u}{:}\sim B)$

6. $\Box(A \supset B)$
$u{:}A$
$\therefore u{:}B$

7. $\Box(A \supset B)$
$\underline{u}{:}A$
$\therefore u{:}B$

8. $\Box(A \supset B)$
$\sim\underline{u}{:}\sim A$
$\therefore \sim\underline{u}{:}\sim B$

9. $\Box(A \supset B)$
$\sim$**u**:B
∴ **u**:$\sim$A

10. $\sim\Diamond(A \bullet B)$
∴ $\sim$(**u**:A • $\sim$**u**:$\sim$B)

13.2b Exercise: LogiCola OB

First appraise intuitively. Then translate into logic and say whether valid (and give a proof) or invalid (no refutation necessary).

1. A logically entails B.
Don’t believe B.
∴ Don’t believe A.

2. You believe A.
∴ You don’t believe not-A. 0298

3. You believe A.
∴ Don’t believe not-A.

4. ∴ If A is self-contradictory, then don’t believe A.

5. ∴ Either believe A or believe not-A.

6. Believe A.
∴ Don’t believe not-A.

7. ∴ Don’t combine believe that A is true with not believing that A is possible.

8. (A and B) entails C.
∴ Don’t combine believing A and believing B and not believing C.

9. A logically entails (B and C).
Don’t believe that B is true.
∴ Believe that A is false.

10. ∴ If A is true, then believe A.

13.3 Believing and willing

Now we'll expand belief logic to cover *willing* as well as *believing*. We'll do this by treating "willing" as *accepting an imperative* – just as we previously treated "believing" as *accepting an indicative*:

u:A = You believe that A
You accept (endorse, assent to, say in your heart) "A is true"

u:$\underline{\mathbf{A}}$ = You will that act A be done
You accept (endorse, assent to, say in your heart) "Let act A be done"

In translating "u:$\underline{\mathbf{A}}$," we'll often use terms more specific than "will" – like "act," "resolve to act," or "desire."[1] Which of these fits depends on whether the imperative is present or future, and whether it applies to oneself or to another. Here are three examples:

If A is present: u:A$\underline{\mathbf{u}}$ = You act (in order) to do A
You accept the imperative for you to do A *now*

If A is future: u:A$\underline{\mathbf{u}}$ = You're resolved to do A
You accept the imperative for you to do A *in the future*

If $u \neq x$: u:A$\underline{\mathbf{x}}$ = You desire (or want) that X do A
You accept the imperative for *X* to do A

And to accept "Would that I had done that" is to wish that you had done it.
0299
There's a subtle difference between "u:A$\underline{\mathbf{u}}$" and "Au":

u:A$\underline{\mathbf{u}}$ = You act (in order) to do A
You say in your heart, "Do A now" (addressed to yourself)

Au = You do A

The first is about what you try or intend to do, while the second is about what you actually do (perhaps accidentally).

Section 12.3 noted that we'd lose important distinctions if we prefixed "O" only to indicatives. Something similar applies here. Consider these three wffs:

[1] "Desire" and similar terms can have a *prima facie* sense ("I have some desire to do A") or an all-things-considered sense ("All things considered, I desire to do A"). Here I intend the latter.

$u:(\exists x)(Kx \bullet R\underline{x})$ = You desire that some who kill *repent*
You say in your heart "Would that some who kill *repent*"

$u:(\exists x)(K\underline{x} \bullet Rx)$ = You desire that some *kill* who repent
You say in your heart "Would that some *kill* who repent"

$u:(\exists x)(K\underline{x} \bullet R\underline{x})$ = You desire that some both *kill and repent*
You say in your heart "Would that some *kill and repent*"

These differ greatly. Underlining shows which parts are desired: repenting, or killing, or killing-and-repenting. If we attached "desire" only to indicative formulas, all three would translate the same, as "You desire that $(\exists x)(Kx \bullet Rx)$" ("You desire that there's someone who both kills and repents"). So "desire" is better symbolized in terms of accepting an imperative.

This imperative formula *tells* you to will something:

$\underline{\mathbf{u}}:\underline{\mathbf{A}}$ = Will that act A be done
Accept (endorse, assent to, say in your heart) "Let act A be done"

Again, our translation can use terms more specific than "will":

If A is present: $\underline{u}:A\underline{u}$ = Act (in order) to do A
Accept the imperative for you to do A *now*

If A is future: $\underline{u}:A\underline{u}$ = Be resolved to do A
Accept the imperative for you to do A *in the future*

If u≠x: $\underline{u}:A\underline{x}$ = Desire (or want) that X do A
Accept the imperative for *X* to do A

Be careful about underlining. Underlining before ":" makes the formula an imperative (instead of an indicative). Underlining after ":" makes the formula about willing (instead of believing). Here are the basic cases: 0300

Indicatives
$u{:}A \quad = \quad$ You believe A.
$u{:}\underline{\mathbf{A}} \quad = \quad$ You will A.
Imperatives
$\underline{\mathbf{u}}{:}A \quad = \quad$ Believe A.
$\underline{\mathbf{u}}{:}\underline{\mathbf{A}} \quad = \quad$ Will A.

These baseball examples may be helpful:

$Hub \quad = \quad$ You hit the ball
$H\underline{\mathbf{u}}b \quad = \quad$ Hit the ball
$OH\underline{\mathbf{u}}b \quad = \quad$ You ought to hit the ball
$RH\underline{\mathbf{u}}b \quad = \quad$ It's all right for you to hit the ball

$u{:}Hub \quad = \quad$ You believe that you'll hit the ball
$u{:}H\underline{\mathbf{u}}b \quad = \quad$ You act (with the intention) to hit the ball
$\underline{\mathbf{u}}{:}Hub \quad = \quad$ Believe that you'll hit the ball
$\underline{\mathbf{u}}{:}H\underline{\mathbf{u}}b \quad = \quad$ Act (with the intention) to hit the ball

13.3a Exercise: LogiCola N (WM & WT)

Translate these English sentences into wffs (use "u" for "you").

Don't act to do A without believing that A would be all right.
$\sim(\underline{\mathbf{u}}{:}A\underline{\mathbf{u}} \bullet \sim\underline{\mathbf{u}}{:}RA\underline{\mathbf{u}})$

1. You want Al to sit down. [Use a for "Al" and Sx for "x sits down."]
2. Believe that Al is sitting down.
3. You believe that Al ought to sit down.
4. Believe that Al intends to sit down.
5. Desire that Al sit down.
6. Eat nothing. [Use Exy for "x eats y."]
7. Resolve to eat nothing.
8. You fall down, but you don't act (in order) to fall down. [Fx]
9. You act to kick the goal, but you don't in fact kick the goal. [Kx]
10. If you believe that you ought to do A, then do A.

11. Don't combine believing that you ought to do A with not acting to do A.

12. Do A, only if you want everyone to do A. (Act only as you'd want everyone to act.) [This is a crude form of Kant's formula of universal law.]

13. If X does A to you, then do A to X. (Treat others as they treat you.) [Use Axy. This principle entails "If X knocks out your eye, then knock out X's eye."]

14. If you do A to X, then X will do A to you. (People will treat you as you treat them.) [This is often confused with the golden rule.]

15. If you want X to do A to you, then do A to X. (Treat others as you want to be treated.) [This is the "literal golden rule."]

16. Don't combine acting in order to do A to X with wanting X not to do A to you.
0301

13.4 Willing proofs

Besides inconsistency in beliefs, there's also inconsistency in will: I might have inconsistent resolutions, violate ends-means consistency, or have moral beliefs that conflict with how I live. Belief logic can generate norms about consistent willing; thus it deals with *practical reason* as well as *theoretical reason*.

Except for having more underlining, proofs with willing formulas work like before. Here's a proof of "Don't combine *believing* that it's wrong for you to do A with *acting* to do A" (these parts clash – since if the *believing* is correct then the *acting* is wrong, and if the *acting* is correct then the *believing* is wrong):

$[\ \therefore \sim(\underline{\mathbf{u}}{:}O{\sim}A\underline{\mathbf{u}} \bullet \underline{\mathbf{u}}{:}A\underline{\mathbf{u}})$ **Valid**
* 1 asm: $(\underline{\mathbf{u}}{:}O{\sim}A\underline{\mathbf{u}} \bullet \underline{\mathbf{u}}{:}A\underline{\mathbf{u}})$
2 $\therefore \underline{\mathbf{u}}{:}O{\sim}A\underline{\mathbf{u}}$ {from 1}
3 $\therefore \underline{\mathbf{u}}{:}A\underline{\mathbf{u}}$ {from 1}
4 $u \therefore O{\sim}A\underline{\mathbf{u}}$ {from 2}
5 $u \therefore A\underline{\mathbf{u}}$ {from 3}
6 $u \therefore {\sim}A\underline{\mathbf{u}}$ {from 4}
7 $\therefore \sim(\underline{\mathbf{u}}{:}O{\sim}A\underline{\mathbf{u}} \bullet \underline{\mathbf{u}}{:}A\underline{\mathbf{u}})$ {from 1; 5 contradicts 6}

The second part of the formula is expressed as "$\underline{\mathbf{u}}{:}A\underline{\mathbf{u}}$" (which is about what you try or intend to do) and not "$A\underline{\mathbf{u}}$" (which is about what you do, perhaps accidentally). The faulty translation "$\sim(\underline{\mathbf{u}}{:}O{\sim}A\underline{\mathbf{u}} \bullet A\underline{\mathbf{u}})$" forbids unintentionally doing what one thinks is wrong; there's no inconsistency in this, except perhaps externally. The correct version forbids this combination: thinking that A is wrong and at the same time acting with the intention of doing A.

13.4a Exercise: LogiCola OW

Say whether valid (and give a proof) or invalid (no refutation necessary).

∴ (u:O∼A<u>u</u> ⊃ ∼<u>u</u>:A<u>u</u>)
[∴ (u:O∼A<u>u</u> ⊃ ∼<u>u</u>:A<u>u</u>) **Invalid** * 1 asm: ∼(u:O∼A<u>u</u> ⊃ ∼<u>u</u>:A<u>u</u>) 2 ∴ u:O∼A<u>u</u> {from 1} 3 ∴ <u>u</u>:A<u>u</u> {from 1} 4 u ∴ A<u>u</u> {from 3} This says: "If you believe it's wrong for you to do A, then don't act to do A"; this leads to problems because it lacks the correct don't-combine form and because your belief may be mistaken. Maybe you believe that it's wrong to treat people fairly; then this formula tells you not to act to treat them fairly.

1. ∴ ∼(<u>u</u>:<u>**A**</u> • <u>u</u>:∼<u>**A**</u>)

2. ∴ <u>u</u>:(B<u>a</u> ⊃ RB<u>a</u>)

3. ∴ (<u>u</u>:B<u>a</u> ∨ <u>u</u>:∼B<u>a</u>)

4. ∴ ∼((<u>u</u>:(A ⊃ <u>**B**</u>) • <u>u</u>:A) • ∼<u>u</u>:<u>**B**</u>)

5. u:(x)OA<u>x</u>
∴ <u>u</u>:A<u>u</u> 0302

6. ∼u:A<u>u</u>
∴ ∼<u>u</u>:OA<u>u</u>

7. ∴ <u>u</u>:(OA<u>u</u> ⊃ A<u>u</u>)

8. ∴ (<u>u</u>:A<u>u</u> ∨ ∼<u>u</u>:OA<u>u</u>)

9. u:A<u>u</u>
∴ ∼<u>u</u>:O∼A<u>u</u>

10. ☐(<u>**A**</u> ⊃ <u>**B**</u>)
∴ ∼(<u>u</u>:O<u>**A**</u> • ∼<u>u</u>:<u>**B**</u>)

13.4b Exercise: LogiCola OW

First appraise intuitively. Then translate into logic and say whether valid (and give a proof) or invalid (no refutation necessary).

1. ∴ Don't combine believing that everyone ought to do A with not acting/resolving to do A yourself. [This is belief logic's version of "Practice what you preach."]

2. ∴ Don't combine resolving to eat nothing with acting to eat this. [Use Exy and t.]

3. "Attain this end" entails "If taking this means is needed to attain this end, then take this means."
∴ Don't combine (1) *wanting* to attain this end and (2) *believing* that taking this means is needed to attain this end and (3) *not acting* to take this means. [Use E for "You attain this end," N for "Taking this means is needed to attain this end," M for "You take this means," and u. The conclusion is an ends-means consistency imperative; you violate it if you want to become a doctor and believe that studying is needed for you to do this and yet you don't act to study.]

4. "Attain this end" entails "If taking this means is needed to attain this end, then take this means."
∴ If you want to attain this end and believe that taking this means is needed to attain this end, then act to take this means. [Use E, N, M, and u. This formulation could tell people with evil ends to do evil things.]

5. ∴ Don't accept "For all x, it's wrong for x to kill," without being resolved that if killing were needed to save your family, then you wouldn't kill. [Kx, N]

6. ∴ Don't accept "For all x, it's wrong for x to kill," without it being the case that if killing were needed to save your family then you wouldn't kill. [Use Kx and N. A draft board challenged a pacifist friend of mine, "If killing were needed to save your family, then would you kill?" My friend answered, "I don't know – I might lose control and kill (it's hard to predict what you'll do in a panic situation); but I now firmly hope and resolve that I wouldn't kill." Maybe my friend didn't satisfy this present formula; but he satisfied the previous one.]

7. ∴ Don't combine accepting "It's wrong for Bob to do A" with wanting Bob to do A.

8. ∴ Don't combine believing that the state ought to execute all murderers with not desiring that if your friend is a murderer then the state execute your friend. [Use s for "the state," Exy for "x executes y," Mx for "x is a murderer," f for "your friend," and u for "you."]

9. ∴ Don't combine acting to do A with not accepting that A is all right.

10. ∴ If you act to do A, then accept that act A is all right.

11. ∴ Don't combine acting to do A with not accepting that A is obligatory. 0303

12. Believe that you ought to do A.
∴ Act to do A.

13. "It's all right for you to do A" entails "It's obligatory that everyone do A."
∴ Don't combine acting to do A with not willing that everyone do A. [The conclusion is a crude version of Kant's formula of universal law. To see that the premise and conclusion are questionable, substitute "become a doctor" for "do A" in both. We'll see a better version of the formula in the next chapter.]

13.5 Rationality translations

Beliefs can be "evident" or "reasonable" for a given person. As I shade my eyes from the bright sun, my belief that it's sunny is *evident*; it's very solidly grounded. As I hear a prediction of rain, my belief that it will rain is *reasonable*; my belief accords with reason but isn't well-grounded enough to be evident. "Evident" expresses a higher certitude than does "reasonable." We'll symbolize these notions as follows:

A is evident to you
= O$\underline{u}$:A
It's obligatory (rationally required) that you believe A
Insofar as intellectual considerations are concerned (including your experiences), you ought to believe A

A is reasonable for you to believe
= R$\underline{u}$:A
It's all right (rationally permissible) that you believe A
Insofar as intellectual considerations are concerned (including your experiences), it would be all right for you to believe A

Neither entails that you believe A; to say that a proposition A that you believe is evident / reasonable, we'll use "(u:A • O**u**:A)" / "(u:A • R**u**:A)." "Evident" and "reasonable" are relative to an individual person; "It's raining" might be evident to someone outside but not to someone inside in a windowless room.

Here are further translations:

It would be unreasonable for you to believe A
= ~R**u**:A
= It's obligatory that you not believe A
= O~**u**:A

It would be reasonable for you to take no position on A
= R(~**u**:A • ~**u**:~A)

It's evident to you that if A then B
= O**u**:(A ⊃ B)

If it's evident to you that A, then it's evident to you that B
= (O**u**:A ⊃ O**u**:B)

You ought not to combine believing A with believing not-A
= O~(**u**:A • **u**:~A)

Since "O" and "R" attach only to imperatives, "Ou:A" and "Ru:A" aren't wffs.

We can almost define "knowledge" simply as "evident true belief": 0304

You know that A
= uKA
= (O**u**:A • (A • u:A))
A is evident to you, A is true, and you believe A

Knowing requires more than just true belief; if you guess right, you have true belief without knowledge. Knowledge must be well-grounded; more than just being *reasonable* (permitted by the evidence), it must be *evident* (required by the evidence). The claim that *knowledge* is *evident true belief* is plausible. But there are cases (like example 10 of §13.6b) where we have one but not the other. So this definition of "knowledge" is flawed; but it's still a useful approximation.

13.5a Exercise: LogiCola N (RM & RT)

Translate these English sentences into wffs. When an example says a belief is evident or reasonable, but doesn't say *to whom*, assume it means evident or reasonable *to you*.

You ought to want Al to sit down.
$\text{O}\underline{\text{u}}\text{:S}\underline{\text{a}}$ We can paraphrase the sentence as "It's obligatory that you say in your heart 'Would that Al sit down.'"

1. You ought to believe that Al is sitting down.

2. It's evident to you that Al is sitting down.

3. It's reasonable for you to believe that Al ought to sit down.

4. Belief in God is reasonable (for you). [G]

5. Belief in God is unreasonable for everyone.

6. It's not reasonable for you to believe that belief in God is unreasonable for everyone.

7. Belief in God is reasonable only if "There is a God" is logically consistent.

8. You ought not to combine believing that there is a God with not believing that "There is a God" is logically consistent.

9. You ought not to combine believing that you ought to do A with not acting to do A.

10. You know that x = x. [Use the flawed definition of knowledge given previously.]

11. If agnosticism is reasonable, then theism isn't evident. [Agnosticism = not believing G and not believing not-G; theism = believing G.]

12. You have a true belief that A. [You believe that A, and it's true that A.]

13. You mistakenly believe A.

14. It would be impossible for you mistakenly to believe A.

15. A is evident to you, if and only if it would be impossible for you mistakenly to believe A. [This idea is attractive but quickly leads to skepticism.]

16. It's logically possible that you have a belief A that's evident to you and yet false.

17. It's evident to all that if they doubt then they exist. [Dx, Ex]

18. If A entails B, and B is unreasonable, then A is unreasonable.

19. It's permissible for you to do A, only if you want everyone to do A.

20. If you want X to do A to you, then you ought to do A to X. [Use Axy. This one and the next are versions of the golden rule.]

21. You ought not to combine acting to do A to X with wanting X not to do A to you. 0305

22. It's necessary that, if you're in pain, then it's evident to you that you're in pain. [Use Px. This claims that "I'm in pain" is a self-justifying belief. Many think that there are two kinds of self-justifying belief: those of experience (as in this example) and those of reason (as in the next example).]

23. It's necessary that, if you believe that x = x, then it's evident to you that x = x. [Perhaps believing "x = x" entails understanding it, and this makes it evident.]

24. If you have no reason to doubt your perceptions and it's evident to you that you believe that you see a red object, then it's evident to you that there is an actual red object. [Use Dx for "x has reason to doubt his or her perceptions," Sx for "x sees a red object," and R for "There is an actual red object." Roderick Chisholm claimed that we need evidential principles like this (but more complex) to show how beliefs about external objects are based on beliefs about perceptions.]

25. If you have no reason to doubt Jenny's sincerity and it's evident to you that she shows pain behavior, then it's evident to you that Jenny feels pain. [Use Bx, Dx, Fx, and j. This exemplifies an evidential principle about knowing other minds.]

13.6 Rationality proofs

Deontic belief proofs, while not requiring further inference rules, often use complex world prefixes like "Du" or "Duu." Here's a proof of a *conscientiousness principle*, "You ought not to combine *believing* that it's wrong for you to do A with *acting* to do A":

[$\therefore O{\sim}(\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$ **Valid**

* 1 asm: $\sim O{\sim}(\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$

* 2 $\therefore R(\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$ {from 1}

* 3 D $\therefore (\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$ {from 2}

4 D $\therefore \underline{u}{:}O{\sim}A\underline{u}$ {from 3}

5 D $\therefore \underline{u}{:}A\underline{u}$ {from 3}

6 Du $\therefore O{\sim}A\underline{u}$ {from 4}

7 Du $\therefore A\underline{u}$ {from 5}

8 Du $\therefore {\sim}A\underline{u}$ {from 6}

9 $\therefore O{\sim}(\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$ {from 1; 7 contradicts 8}

We get to line 5 using propositional and deontic rules. Lines 6 and 7 follow using rule B+. Here we write belief world prefix "u" after the deontic world prefix "D" used in lines 3 to 5; world Du is a belief world of u that depends on what deontic world D tells u to accept. We soon get a contradiction.

"$O{\sim}(\underline{u}{:}O{\sim}A\underline{u} \bullet \underline{u}{:}A\underline{u})$" is a **formal ethical principle** – an ethical principle that can be formulated using the abstract notions of our logical systems plus variables (like "u" and "A") that stand for any person and action. The next chapter will focus on another formal ethical principle – the golden rule. 0306

13.6a Exercise: LogiCola O (R & M)

Say whether valid (and give a proof) or invalid (no refutation necessary).

$R\underline{u}{:}O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$
$\therefore R\underline{u}{:}O\underline{\mathbf{A}}$

	1	$R\underline{u}{:}O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$	**Valid**
		$[\,\therefore R\underline{u}{:}O\underline{\mathbf{A}}$	
*	2	asm: $\sim R\underline{u}{:}O\underline{\mathbf{A}}$	
	3	$\therefore O\sim\underline{u}{:}O\underline{\mathbf{A}}$	{from 2}
	4	$D \therefore \underline{u}{:}O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$	{from 1}
*	5	$D \therefore \sim\underline{u}{:}O\underline{\mathbf{A}}$	{from 3}
*	6	$Du \therefore \sim O\underline{\mathbf{A}}$	{from 5}
	7	$Du \therefore O(\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$	{from 4}
*	8	$Du \therefore R\sim\underline{\mathbf{A}}$	{from 6}
	9	$DuDD \therefore \sim\underline{\mathbf{A}}$	{from 8}
	10	$DuDD \therefore (\underline{\mathbf{A}} \bullet \underline{\mathbf{B}})$	{from 7}
	11	$DuDD \therefore \underline{\mathbf{A}}$	{from 10}
	12	$\therefore R\underline{u}{:}O\underline{\mathbf{A}}$	{from 2; 9 contradicts 11}

(If you can follow this example, you needn't fear proofs involving complex world prefixes.)

1. $\Box(A \supset B)$
$\sim R\underline{u}{:}B$
$\therefore \sim R\underline{u}{:}A$

2. $O\sim\underline{u}{:}A$
$\therefore O\underline{u}{:}\sim A$

3. $R(\sim\underline{u}{:}A \bullet \sim\underline{u}{:}\sim A)$
$\therefore \sim O\underline{u}{:}A$

4. $R\underline{u}{:}\sim A$
$\therefore R\sim\underline{u}{:}A$

5. $O\underline{a}{:}(C \bullet D)$
$\therefore O\underline{b}{:}C$

6. $\therefore O\sim(\underline{u}{:}A \bullet \sim\underline{u}{:}\Diamond A)$

7. $\therefore (R\underline{u}{:}A \supset \Diamond A)$

8. ☐(A ⊃ B)
∴ (R~<u>u</u>:B ⊃ R<u>u</u>:~A)

9. R<u>u</u>:OA<u>u</u>
∴ R<u>u</u>:◇Au

10. O<u>u</u>:(A ⊃ OB<u>u</u>)
∴ ~(<u>u</u>:A • ~<u>u</u>:B<u>u</u>)

13.6b Exercise: LogiCola O (R & M)

First appraise intuitively. Then translate into logic and say whether valid (and give a proof) or invalid (no refutation necessary). Use G for "There is a God" and u for "you." When an example says a belief is evident or reasonable, but don't say *to whom*, assume it means evident or reasonable *to you*.

1. Theism is evident.
∴ Atheism is unreasonable.   [Theism = believing G; atheism = believing not-G.]

2. Theism isn't evident.
∴ Atheism is reasonable.

3. ∴ You ought not to combine *believing* you ought to do A with *not acting* to do A.

4. ∴ If you believe you ought to do A, then you ought to do A.

5. "All men are endowed by their creator with certain unalienable rights" is evident.
"All men are endowed by their creator with certain unalienable rights" entails "There is a creator."
∴ "There is a creator" is evident.   [Use E and C. The opening lines of the US Declaration of Independence claim E to be self-evident.] 0307

6. It would be reasonable for you to believe that A is true.
It would be reasonable for you to believe that B is true.
∴ It would be reasonable for you to believe that A and B are both true.

7. "If I'm hallucinating, then physical objects aren't as they appear to me" is evident to me.
It's not evident to me that I'm not hallucinating.
∴ It's not evident to me that physical objects are as they appear to me.   [Use H, P, and i. This argument for skepticism is essentially from Descartes.]

8. “If I’m hallucinating, then physical objects aren’t as they appear to me” is evident to me.
If I have no special reason to doubt my perceptions, then it’s evident to me that physical objects are as they appear to me.
I have no special reason to doubt my perceptions.
∴ It’s evident to me that I’m not hallucinating. [Use H, P, D, and i. This is John Pollock’s answer to the previous argument.]

9. It’s evident to you that taking this means is needed to attain this end.
“Attain this end” entails “If taking this means is needed to attain this end, then take this means.”
∴ You ought not to combine wanting to attain this end with not acting to take this means. [Use N for “Taking this means is needed to attain this end,” E for “You attain this end,” M for “You take this means,” and u.]

10. Al believes that Smith owns a Ford.
It’s evident to Al that Smith owns a Ford.
Smith doesn’t own a Ford.
Smith owns a Chevy.
Al believes that Smith owns a Ford or a Chevy.
Al doesn’t know that Smith owns a Ford or a Chevy.
∴ Al has an evident true belief that Smith owns a Ford or a Chevy; but Al doesn’t know that Smith owns a Ford or a Chevy. [Use a for “Al,” F for “Smith owns a Ford,” C for “Smith owns a Chevy,” and K for “Al knows that Smith owns a Ford or a Chevy.” This argument from Edmund Gettier attacks the definition of *knowledge* as *evident true belief*.]

11. It’s evident to you that if it’s all right for you to hit Al then it’s all right for Al to hit you.
∴ Don’t combine acting to hit Al with believing that it would be wrong for Al to hit you. [Use Hxy, u, and a. The premise is normally true; but it could be false if you and Al are in different situations (maybe Al needs to be hit to dislodge food he’s choking on). The conclusion resembles the golden rule.]

12. ∴ It’s reasonable to want A to be done, only if it’s reasonable to believe that A would be all right.

13. It’s evident that A is true.
∴ A is true. 0308

14. It's reasonable to combine believing that there is a perfect God with believing T.
T entails that there's evil in the world.
∴ It's reasonable to combine believing that there is a perfect God with believing that there's evil in the world. [Use G, T, and E. Here T (for "theodicy") is a reasonable explanation of why God permits evil, perhaps "The world has evil because God, who is perfect, wants us to make significant free choices to struggle to bring a half-completed world toward its fulfillment; moral evil comes from the abuse of human freedom and physical evil from the half-completed state of the world."]

15. It's evident to you that if there are moral obligations then there's free will.
∴ Don't combine accepting that there are moral obligations with not accepting that there's free will. [M, F]

16. Theism is reasonable.
∴ Atheism is unreasonable.

17. Theism is evident.
∴ Agnosticism is unreasonable. [Agnosticism = not believing G and not believing not-G.]

18. ∴ It's reasonable for you to believe that God exists, only if "God exists" is consistent. [Belief logic regards a belief as "reasonable" only if *in fact* it's consistent. In a more subjective sense, someone could "reasonably" believe a proposition that's reasonably but incorrectly taken to be consistent.]

19. ∴ If A is unreasonable, then don't believe A.

20. You ought not to combine accepting A with not accepting B.
∴ If you accept A, then accept B.

21. ∴ You ought not to combine wanting A not to be done with believing that A would be all right.

22. It's reasonable not to believe that there is an external world.
∴ It's reasonable to believe that there's no external world. [E]

23. It's reasonable to believe that A ought to be done.
∴ It's reasonable to want A to be done.

24. ∴ Either theism is reasonable or atheism is reasonable.

25. It's evident to you that if the phone is ringing then you ought to answer it.
It's evident to you that the phone is ringing.
∴ Act on the imperative "Answer the phone." [P, Ax]

26. A entails B.
Believing A would be reasonable.
∴ Believing B would be reasonable.

27. Atheism isn't evident.
∴ Theism is reasonable. 0309

28. Atheism is unreasonable.
Agnosticism is unreasonable.
∴ Theism is evident.

29. A entails B.
You accept A.
It's unreasonable for you to accept B.
∴ Don't accept A, and don't accept B.

30. It would be reasonable for anyone to believe A.
∴ It would be reasonable for everyone to believe A. [Imagine a controversial issue where everyone has the same evidence. Could it be more reasonable for the community to disagree? If so, the premises of this argument might be true but the conclusion false.]

13.7 A sophisticated system

The system of belief logic that we've developed is oversimplified in three ways. We'll now sketch a more sophisticated system.

First, "One ought to be consistent" requires qualification. For the most part, we do have a duty to be consistent. But, since "ought" implies "can," this duty is nullified when we're unable to be consistent; such inability can come from emotional turmoil[1] or our incapacity to grasp complex logical relations. And the obligation to be consistent can be overridden by other factors; if Dr Evil would destroy the world unless we were inconsistent in some respect, then surely our duty to be consistent would be overridden. And the duty to be consistent applies, when it does, only to persons; yet our principles so far

[1] Perhaps you see (and believe) that your wife was in a car that blew up and you believe that anyone in such a car would be dead – but you're psychologically unable at the moment to believe that your wife is dead. Then you're psychologically unable at the moment to be consistent about this.

would entail that rocks and trees also have a duty to be consistent.

For these reasons, it would be better to qualify our "You ought to be consistent" principle, as in the following rough formulation:[1]

> If you are a person able to be consistent in certain ways, grasp (or should grasp) the logical relationships, and your being consistent wouldn't have disastrous consequences, then you ought to be consistent in these ways.

Let's abbreviate the qualification in the box ("You are …") as "Qu." Then we could reformulate our inference rules by adding a "Qu" premise:

B+
$\underline{\mathbf{u}}$:A, Qu → u ∴ A, use any string of u's

B−
$\sim\underline{\mathbf{u}}$:A, Qu → u ∴ ~A, use a *new* string of u's

0310 With these changes, we'd need plentiful "Qu" provisos in the previous sections.

A second problem is that our system can prove a *conjunctivity principle*:

> $O\sim((\underline{\mathbf{u}}{:}A \bullet \underline{\mathbf{u}}{:}B) \bullet \sim\underline{\mathbf{u}}{:}(A \bullet B))$
> You ought not to combine believing A and believing B and not believing A-and-B

This leads to questionable results in the *lottery paradox*. Suppose six people have an equal chance to win a lottery. You know that one of the six will win; but the probability is against any given person winning. Presumably it could be reasonable for you to accept statements 1 to 6 without also accepting statement 7 (which means "None of the six will win"):

1. Person 1 won't win.
2. Person 2 won't win.
3. Person 3 won't win.
4. Person 4 won't win.
5. Person 5 won't win.

[1] Section 2.3 of my *Formal Ethics* (London: Routledge, 1996) has additional qualifications.

6. Person 6 won't win.

7. Person 1 won't win, person 2 won't win, person 3 won't win, person 4 won't win, person 5 won't win, and person 6 won't win.

But multiple uses of our conjunctivity principle would entail that one ought not to accept statements 1 to 6 without also accepting their conjunction 7. So the conjunctivity principle, which is provable using our rules B+ and B–, sometimes leads to questionable results.

I'm not completely convinced that it's reasonable to accept statements 1 to 6 but not accept 7. If it *is* reasonable, then we have to reject the conjunctivity principle and modify our consistency ideal. Let's call the ideal of "completely consistent" defined in §13.2 *broad consistency*. Perhaps we should strive, not for this, but for *narrow consistency*. Let S be the set of indicatives and imperatives that X accepts; then X is *narrowly consistent* if and only if:

1. every pair of items of set S is logically consistent, and
2. X accepts whatever follows from any single item of set S.

Believing the six lottery statements but not their conjunction is narrowly consistent but not broadly consistent.

To have our rules mirror the ideal of narrow consistency, we'd add to rules B+ and B– that any belief world prefix used in these rules cannot have occurred more than once in earlier lines. With this change, only a few arguments in this chapter would cease being provable. And many of these could be salvaged by adding an additional conjunctivity premise like the following (which would be true in many cases): "You ought not to combine believing A and believing B and not believing A-and-B." Conjunctivity presumably fails only in rare lottery-type cases.

The third problem is that we've been translating these two statements the same way, as "O**u**:A," even though they don't mean the same thing: 0311

"You ought to believe A" $\neq$ "A is evident to you"

Suppose you ought to trust your wife and give her the benefit of every reasonable doubt; you *ought to believe* what she says, even though the evidence isn't so strong as to make this belief *evident*. Here there's a difference between "ought to believe" and "evident." And so it may be better to use a different symbol (perhaps "O*") for "evident":

You ought to believe A
= O**u**:A
All things considered, you ought to believe A

A is evident to you
= O***u**:A
Insofar as intellectual considerations are concerned (including your experiences), you ought to believe A

"O" is an all-things-considered "ought," while "O*" is a *prima facie* "ought" that considers only the intellectual basis for the belief. If we added "O*" to our system, we'd need corresponding deontic inference rules for it. Since "O***A**" is a *prima facie* "ought," it wouldn't entail the corresponding imperative or commit one to action; so we'd have to weaken the rule for dropping "O*" so we couldn't derive "**u**:A" from "O***u**:A."

These three refinements would overcome some problems but make our system much harder to use. We seldom need the refinements. So we'll keep the naïve belief logic of earlier sections as our "official system" and build on it in the next chapter. But we'll be conscious that this system is oversimplified in various ways. If and when the naïve system gives questionable results, we can appeal to the sophisticated system to clear things up.

14 A Formalized Ethical Theory

This chapter gives a precise logical formulation of an ethical theory, one that builds on ideas from Immanuel Kant and R. M. Hare.[1] This gives an example of how to use logical systems to formalize larger philosophical views. As in the belief-logic chapter, we'll systematize *consistency norms*. But here we feature the golden rule (roughly, "Treat others as you want to be treated").

We'll first consider practical reason, highlighting the role of consistency. Then we'll focus on the golden rule. After seeing problems with the usual wording, we'll give a better formulation and an intuitive argument for it. Then we'll add symbols and inference rules to formalize these ideas. We'll end with a formal proof of the golden rule in logical symbols.

14.1 Practical reason

The most important elements of practical reason are *factual understanding*, *imagination*, and *consistency*. As we decide how to act on important matters, and as we form related desires or moral beliefs, we ought as far as practically possible to be vividly aware of the relevant facts, avoid falsehoods, and be consistent.

Factual understanding requires that we know the facts of the case: circumstances, alternatives, consequences, and so on. To the extent that we're misinformed or ignorant, our moral thinking is flawed. Of course, we can never know *all* the facts; and often we have no time to research a problem and must act quickly. But we can act out of greater or lesser knowledge. Other things being equal, a more informed judgment is a more rational one.

We also need to understand ourselves, and how our feelings and moral beliefs originated; we can to some extent neutralize our biases if we understand their origin. Some people are hostile toward a group because they were brought up that way, especially through false stereotypes. Their attitudes might change if they understood the source of their hostility and broadened

[1] For fuller accounts, see my *Ethics and the Golden Rule* (New York: Routledge, 2013) and *Ethics and Religion* (New York: Cambridge University Press, 2016), my technical *Formal Ethics* (New York: Routledge, 1996), or my simpler *Ethics: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2018). See also Immanuel Kant's *Groundwork of the Metaphysics of Morals* (New York: Harper & Row, 1964) and R. M. Hare's *Freedom and Reason* (New York: Oxford University Press, 1963).

their experience 0313 and knowledge; if so, then their attitudes are less rational, since they exist because of a lack of experience and self-knowledge.

Imagination (role reversal) is a vivid and accurate awareness of what it would be like to be in the place of those affected by our actions. This differs from just knowing facts. So in dealing with poor people, besides knowing facts about them, we also need to appreciate and envision what these facts mean to their lives; movies, literature, and personal experience can help us to visualize another's life. We also need to appreciate future consequences of our actions on ourselves; knowing that drugs would have harmful effects on us differs from being able to imagine these effects in a vivid and accurate way.

Consistency (which we'll explore in the next section) demands, among other things, a coherence between our beliefs, our ends and means, and our moral judgments and how we live. We need all the dimensions of moral rationality working together for our practical thinking to be fully reasonable; consistency is important but isn't everything. Appeals to consistency in ethics are often distressingly vague; my goal here is to clarify and defend consistency norms.

The most important part of practical reason is the golden rule. As we decide how to act toward others, we ought as far as practically possible to be vividly aware of the relevant facts (especially about how our action affects the other person and what it would be like to be treated that way), avoid falsehoods (about this), and be consistent (so we don't treat another as we're unwilling that we be treated in the same situation).

14.2 Consistency

Consistency is the basis for key elements of practical reason, including reflective equilibrium, ends-means rationality, and the golden rule. Our belief-logic chapter touched on these three consistency norms:[1]

- Logicality: Avoid inconsistency in beliefs.
- Ends-means consistency: Keep your means in harmony with your ends.
- Conscientiousness: Keep your actions, resolutions, and desires in harmony with your moral beliefs.

Our belief logic contains *logicality* norms forbidding inconsistent beliefs:

[1] We noted at the end of the last chapter that consistency duties require qualifiers, like "insofar as you're able to be consistent in these ways and no disaster would result from so doing ...." This also applies to the golden rule. We'll regard such qualifiers as implicit throughout.

$(\sim\Diamond(A \bullet B) \supset \sim(\underline{\mathbf{u}}{:}A \bullet \underline{\mathbf{u}}{:}B))$
Don’t combine inconsistent beliefs.
If A is inconsistent with B, then don’t combine *believing A* with *believing B*. 0314

$(\Box(A \supset B) \supset \sim(\underline{\mathbf{u}}{:}A \bullet \sim\underline{\mathbf{u}}{:}B))$
Don’t believe something without believing what follows from it.
If A logically entails B, then don’t combine *believing A* with *not believing B*.

Consistency pushes us toward a *reflective equilibrium* in our thinking between principles and concrete judgments. Suppose I accept an appealing moral principle but reject an unappealing concrete judgment that it logically entails. Then something has to give; I have to reject the principle or accept the concrete judgment. Before deciding which to do, I need to investigate the principle further. Much moral thinking follows this reflective-equilibrium pattern.

Our belief logic can prove this *ends-means consistency* argument (§13.4b #3):

$\Box(\underline{\mathbf{E}} \supset (N \supset \underline{\mathbf{M}}))$
$\therefore \sim((\underline{\mathbf{u}}{:}\underline{\mathbf{E}} \bullet \underline{\mathbf{u}}{:}N) \bullet \sim\underline{\mathbf{u}}{:}\underline{\mathbf{M}})$

“Attain this end” entails “If taking this means is needed to attain this end, then take this means.”
∴ Don’t combine *wanting* to attain this end, *believing* that taking this means is needed to attain this end, and *not acting* to take this means.[1]

Ends and means are important to human life. We have many goals – including food, shelter, health, companionship, and meaningfulness. Practical reason has us try to understand our goals, investigate how to satisfy them, satisfy ends-means consistency, and reject ends or means that lead us to violate golden-rule consistency.

Our belief logic also can prove *conscientiousness* principles that prescribe a harmony between our moral beliefs and how we live. Here’s an example:

$\sim(\underline{\mathbf{u}}{:}OA\underline{\mathbf{u}} \bullet \sim\underline{\mathbf{u}}{:}A\underline{\mathbf{u}})$
Don’t combine *believing* that you ought to do A with *not acting* to do A.

This is a **formal ethical principle** – an ethical principle that can be formulated

[1] If we added “[c]” for causal necessity (see Arthur Burks’s *Chance, Cause, Reason* (Chicago: University of Chicago Press, 1977)) to our system, then “$\sim((\underline{u}{:}\underline{E} \bullet \underline{u}{:}[c](\sim M \supset \sim E)) \bullet \sim\underline{u}{:}\underline{M})$” could be provable by itself: “Don’t combine *wanting* to attain this end, *believing* that taking this means is needed to attain this end, and *not acting* to take this means.”

using the abstract notions of our logical systems plus variables (like "u" and "A," which stand for any person and action). All our consistency requirements are *formal* in this sense.

Consistency is important in criticizing norms. Suppose I was taught to discriminate against short people and to believe *shortism*: "All short people ought to be beat up, just because they're short." Now shortism entails "If I were short, then I ought to be beat up"; so consistency in beliefs commits me to accepting this too. But then, by conscientiousness, I'm committed to *desiring that if I were short then I be beat up*. So consistency forbids this combination: 0315

- I believe "All short people ought to be beat up, just because they're short."
- I don't desire that if I were short then I be beat up.

When I understand short people (including how it feels for them to be beat up) and how my negative attitudes about them originated (through social indoctrination), and when I vividly imagine myself being beaten up in their place, then I likely won't desire that if I were short then I be beat up. But then I'm inconsistent in accepting shortism. The same general approach – which may remind us of the golden rule – can be used to counter other discriminatory principles (racial, religious, gender, sexual orientation, etc.).

Here are three further formal consistency requirements:

- Impartiality: Make similar evaluations about similar actions, regardless of the individuals involved.[1]
- Golden rule: Treat others only as you consent to being treated in the same situation.
- Formula of universal law: Act only as you're willing for anyone to act in the same situation – regardless of imagined variations of time or person.

We'll add logical machinery for all three, but mostly focus on the golden rule.

14.3 The golden rule

GR says "Treat others as you want to be treated." GR is a global standard, endorsed by nearly every religion and culture, important for professionals and families across the planet, and a key part of a growing global-ethics

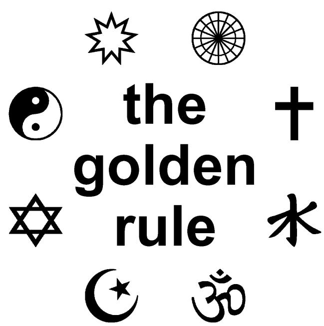
[1] I defend only a weak impartiality, not a strong utilitarian impartiality that claims that we ought to promote everyone's good equally. Weak impartiality lets us accept that we ought to have greater concern for our children than for strangers, so long as we accept that other parents also ought in similar cases to have a greater concern for their children.

movement.

Here's a story to introduce GR.[1] There once was a grandpa who lived with his family. As Grandpa grew older, he began to slobber and spill his food; so the family had him eat alone. When he dropped his bowl and broke it, they scolded him and got him a cheap wooden bowl. Grandpa was so unhappy. Now one day the young grandson was working with wood. "What are you doing?" Mom and Dad asked. "I'm making a wooden bowl," he said, "for when you two get old and must eat alone." Mom and Dad then looked sad and realized how they were mistreating Grandpa. So they decided to keep quiet when he spills his food and let him eat with the family.

The heart of the golden rule is *switching places*. You step into another's shoes. What you do to Grandpa, you imagine being done to you. You ask, "Am I willing that if I were in the same situation then I be treated that same way?"
0316

The golden rule seems simple. But the usual loose wordings invite objections; many academics dismiss GR as a folksy proverb that self-destructs when analyzed carefully. But I think that we just need to understand GR more clearly. I put my improved wording on a shirt.[2] It has "the golden rule" with symbols for eight GR religions (Bahá'í, Buddhism, Christianity, Confucianism, Hinduism, Islam, Judaism, and Taoism). It also has my GR formula:



Treat others only as you consent to being treated in the same situation.

My formula is intended to help us apply GR to difficult cases.

My GR formula commands consistency. It demands a fit between my act toward another and my desire about how I'd be treated in the same situation. It doesn't replace other moral norms or theories, or give all the answers. It

[1] This traditional "The old man and his grandson" story was published by the Brothers Grimm in 1812 and is online (http://www.gutenberg.org/ebooks/2591).

[2] You can get your own golden-rule shirt, in many styles and colors, from my http://www.harryhiker.com/gr GR Web page. This popular page also has GR information, videos, stories, chronology, links, and so on.

doesn’t say specifically what to do (so it doesn’t command bad actions if we have flawed desires); instead, it forbids an inconsistent combination:

- I do A to another.
- I’m unwilling that if I were in the same situation then A be done to me.[1]

GR, far from being vague, is a precise consistency test. Suppose I force Grandpa to eat alone. I switch places in my mind: I imagine that *I* am forced to eat alone in the same situation. Do I condemn this same act done to me? Then I condemn how I treat Grandpa. *I condemn how I treat another, if I condemn the same act when I imagine it done to me in the same situation*.

People who reject GR usually understand it crudely, often as:

Literal GR: If you want X to do A to you, then do A to X.

The literal GR “$(u{:}A\underline{\mathbf{x}}u \supset A\underline{\mathbf{u}}x)$” has no same-situation clause and it commands a specific act (instead of forbidding an inconsistent combination). This literal GR often works well. Suppose you want Lucy to be kind to you; then you’re to be kind to her. Or suppose you want Adam not to hurt you (or rob you, or be rude to you); then you aren’t to do these things to him. These applications seem sensible. But the literal GR can lead to absurdities in two ways. 0317

First, you may be in a *different situation* from the other person. Consider this instance of the literal GR:

Suppose your father is hard of hearing: If you want your father not to speak more loudly to you (who hear well), then don’t speak more loudly to him.

This ignores differences between you and your father. To get around the problem, you need a same-situation qualifier: “How do I desire that I’d be treated *if I were in the same situation as my father* (and thus hard of hearing)?” You desire that if you were in his same situation then people would speak loudly to you; so you speak loudly to him.

We can take “same” situation here as “exactly similar” or “relevantly similar.” In the first case, I imagine myself in my father’s *exact place* (with all his properties). In the second, I imagine myself having those properties of my father (such as being hard of hearing) that I think are or might be *relevant* to deciding how to speak to him. Either approach works fine.

[1] “Unwilling” here can be taken as “objecting to.” Then the forbidden combination is: (1) I do A to another and (2) I object to the idea of A being done to me in the same situation. If we’re playing chess, I object to the idea of your *cheating* to beat me (I’m unwilling that you do this) but I don’t object to the idea of your beating me if you do so fairly (I’m in this sense “willing” that you do this). (I thank Tom Carson for this clarification and example.)

Here's another case where the literal GR leads to problems:

> To a patient: If you want the doctor to remove your appendix, then remove the doctor's appendix.

Again, we need a same-situation qualifier. The patient clearly doesn't desire that if he were in the place of his doctor (with a healthy appendix), then his appendix be removed by a sick patient ignorant of medicine. As you apply GR, ask this:

> Am I willing that if I were **in the same situation** then this be done to me?

The other person's situation includes likes and dislikes. So if you're a waiter who hates broccoli, but your customer likes and orders it, then you imagine being served broccoli in a hypothetical situation where you like and order it.

GR is about *our present reaction to a hypothetical situation*; it isn't about how we'd react if we were in that situation. Suppose I have a two-year-old son, little Will, who puts fingers into electrical outlets. I try to discourage him from doing this, but nothing works. Finally, I decide that I need to punish him when he does it. I want to see if I can punish him without violating GR. I should ask this:

> Am I **willing that if** I were in the same situation as little Will then I be punished?

I'd answer yes (since punishment would likely have saved my life). I might add, "I'm thankful that my parents punished me in such cases, even though I wasn't pleased then." So here I can punish my child without breaking GR, since I'm **willing that if** I were in the same situation then I be treated the same way. 0318

I've been underlining "**willing that if**," because this phrase guards against a common GR misunderstanding, one that would force us to do whatever the other person wants. People often ask, wrongly, "If I were in the other person's place, how would I then want to be treated?" Now if you were in little Will's place (not knowing about electricity and not wanting to be punished), then you wouldn't want to be punished. Misapplying GR, we'd conclude that we shouldn't punish Will for putting his fingers into outlets. So it's better to apply GR as explained above. I can punish little Will (to save his life), since I'm now **willing that if** I were in his situation then I be punished. In asking the GR question, say "**willing that if**":

Am I **willing that if** I were in the same situation then this be done to me?

Immanuel Kant's 1785 objection to GR rests on this confusion. Here you're a judge, about to sentence a dangerous criminal to jail. The criminal protests and appeals (incorrectly) to GR: "If you were in my place, you wouldn't want to be sent to jail; so by the golden rule you can't send me to jail." You should respond: "I can send you to jail, because I'm now **willing that if** I were in your place (as a dangerous criminal) then I be sent to jail." You could add, "If I do such things, then please send me to jail too!"[1]

Sometimes we need to act against what others want. We may need to stop a baby who wants to put fingers into outlets, refuse a salesperson who wants to sell us overpriced products, fail a student who doesn't work, defend ourselves against an attacker, or jail a dangerous criminal. GR lets us act against what others want, as long as we're now **willing that if** we were in their situation then we be treated similarly.

Recall that the literal GR can lead to absurdities in two ways. We dealt with the *different-circumstances* problem by adding a same-situation clause. A second problem is that the literal GR can tell us to do bad things if we have *flawed desires* about how we're to be treated. I'll give three examples.

There once was a woman named Electra. Electra wanted to follow GR, but she got her facts wrong; she thought electrical shocks were pleasant. Since she wanted others to shock her, she applied the literal GR and shocked them:

> To Electra (who thinks electrical shocks are pleasant): If you want others to give you electrical shocks, then give them electrical shocks.

Given flawed desires, the literal GR can command evil actions. 0319

We'll use a triple strategy for dealing with flawed desires. (1) Emphasize that GR, instead of telling us specifically what to do, just forbids a combination:

- I give electrical shocks to another.
- I'm unwilling that if I were in the same situation then electrical shocks be given to me.

Since the consistency GR doesn't say specifically what to do, it doesn't tell Electra to do evil things (like shock others).

(2) Emphasize that GR consistency, to lead reliably to right action, needs to combine with other things, like knowledge and imagination. If we're misin-

[1] *Groundwork of the Metaphysics of Morals*, trans. H. Paton (New York: Harper & Row, 1964), p. 97 footnote. GR requires wide scope, roughly, "I desire that if A happened then B be done" "$i{:}(A \supset \underline{B})$," instead of "If A happened then I'd desire that B be done" "$(A \supset i{:}\underline{B})$."

formed, then we might do evil things without violating GR consistency. Here Electra shocks others (an evil thing) but satisfies GR consistency (she's willing that she be shocked in similar cases), since she's misinformed and thinks these shocks are pleasurable.

(3) Use reason against flawed desires. Here we'd show Electra that electrical shocks are painful (perhaps by giving her a small one). Once she understands this, GR consistency will lead her away from shocking others.

Here's another example. Mona hates herself and wants others to hate her; so, following the literal GR, she hates others. (1) Again, the correctly formulated GR just forbids a combination and so doesn't prescribe that she hate others. (2) GR consistency, to lead reliably to right action, needs to combine with other things (like knowledge, imagination, and here *a healthy self-love*). (3) We can use reason against Mona's flawed desires. We can try to help Mona understand *why* she hates herself and *how* to neutralize this hatred – by not fixating on her negatives, by seeing herself and her good points in a more balanced way, and, if she's a believer, by appreciating that God loves her. Once Mona regains a healthy self-love, GR consistency will lead her more readily to love others.

Or suppose Adolph is a Nazi who so hates Jews that he kills them and desires that he be killed if he were Jewish (or found to be Jewish). The literal GR would tell Adolph to kill Jews:

> To Adolph (a Jew-hating Nazi): If you want others to kill you if you were Jewish, then kill others who are Jewish.

Again, we can make three points. (1) GR, properly formulated, doesn't command specific actions but instead just forbids an inconsistent combination:

- I kill others just because they're Jewish.
- I'm unwilling that if I were Jewish then I'd be killed just because I'm Jewish.

Since the consistency GR doesn't say specifically what to do, it doesn't tell Adolph to kill Jews. (2) GR consistency, to lead reliably to right action, has to combine with other things (like knowledge, imagination, and here *rational desires*). (3) We can use reason against Adolph's flawed desires. We can try to 0320 help him understand *why* he hates Jews so much, even desiring that he be killed if he were found out to be Jewish. His anti-Jewish hatred likely has its source in things that can be rationally criticized. Maybe Adolph thinks Aryans are superior to Jews and racially pure; we can criticize this on factual grounds. Or maybe Adolph was taught to hate Jews by his family and friends; maybe they hated Jews, called them names, and spread false stereotypes about them. And so his anti-Jewish desires likely came from false beliefs and social conditioning; such flawed desires would diminish if he understood their origin and broadened his experience and knowledge of Jews in an open and personal way. With greater rationality, Adolph wouldn't desire that he'd be killed if found out to be Jewish – and GR would be a powerful tool against his racism.

While this example was about a Nazi, the same idea applies to those who desire that they be mistreated if they were black, female, gay, or whatever. Such desires are likely flawed (as based on a social conditioning that uses false beliefs and stereotypes) and would be given up if we expanded our knowledge and experience of the group in an open and personal way.

As you apply the golden rule, keep in mind that it doesn't work alone. **KITA** (Know-Imagine-Test-Act) is an acronym to help us remember some key elements for using GR wisely:

KITA: Know Imagine Test Act

Know: "How would my action affect others?"

Imagine: "What would it be like to have this done to me in the same situation?"

Test for consistency: "Am I now **willing that if** I were in the same situation then this be done to me?"

Act toward others only as you're willing to be treated in the same situation.

To lead reliably to right action, GR consistency needs to build on things like knowledge, imagination, creativity, rationalized desires, and a healthy self-love.

GR can fit many perspectives. Philosophically, GR could be a self-evident truth (or derivable from such), God's will, a cultural convention, a social contract for mutual advantage, socially useful, reflecting our feelings, promoting self- interest (since it brings self-respect and better treatment from others), and so on. Religiously, GR is part of Bahá'í, Buddhism, Christianity, Confucianism, Hinduism, Islam, Judaism, Sikhism, Taoism, Zoroastrianism, and so on. Diverse groups share GR. The golden rule can be a point of unity in a diverse world.

14.4 Starting the GR proof

What sort of *inconsistency* do we have when we violate the golden rule? Clearly we don't have an inconsistency between beliefs; what clashes here isn't beliefs 0321 but rather actions and desires. But why is it inconsistent to violate GR?

Consistency in a broad sense includes things like ends-means consistency, conscientiousness, and impartiality. GR follows from conscientiousness and impartiality. Suppose that you're *conscientious* and *impartial* in the required senses, and yet you want to steal Detra's bicycle. Being conscientious, you won't steal her bicycle unless you think this act is all right (permissible). Being impartial, you won't think this act is all right unless you think that if you were in the same situation then it would be all right for your bike to be stolen. Being conscientiousness, you won't think this unless you're willing that if you were in the same situation then your bike be stolen. So if you're conscientious and impartial, then you won't steal Detra's bicycle unless you're willing that your bike be stolen in the same situation. Here's a diagram:

You steal Detra's bicycle

then if you're conscientious ⇒

You believe it would be all right for you to steal her bicycle

then if you're impartial ⇒

You believe that if you were in the same situation then it would be all right for your bicycle to be stolen

then if you're conscientious ⇒

You're willing that if you were in the same situation then your bicycle be stolen

So if we're conscientious and impartial, then we'll follow GR: we won't do something to another unless we're willing that it be done to us in the same situation. If we violate GR, then we violate either conscientiousness or impartiality or both. So if we assume that we ought to be conscientious and impartial, then we can *deduce* that we ought to follow the golden rule.

So my GR can be based on an abstract argument; similar reasoning justifies many variations. We can consider someone else we care about (maybe our daughter) on the receiving end of the action. We can give consistency conditions, not for *doing something*, but for *wanting something to be done* or for *holding a moral belief*. A *multi-party* GR has us satisfy GR toward each affected party. A *future-regard form* has us imagine ourselves suffering the future consequences of our present action. A *self-regard form* has us imagine someone we care about doing the self-destructive thing we're doing to ourselves. My formula of universal law is a generalized GR that contains many of these: "Act only as you're willing for anyone to act in the same situation, regardless of where or when you imagine yourself or others."

So GR follows from the requirements to be conscientious and impartial. But why be conscientious and impartial? Why care about consistency at all?

Different views could answer differently. Maybe we ought to be consistent because this is inherently right; our minds see consistency as the first duty of a 0322 rational being. Or maybe we accept consistency because it's commanded by God, useful to social life, accords with how we want to live, or promotes our self-interest (since inconsistency brings lowered self-respect, painful "cognitive dissonance," and social sanctions). I'll abstract from such issues here and assume only that there's *some* strong reason to be consistent, in a broad sense that includes being conscientious and impartial. I won't worry about the details. I'm trying to develop consistency norms that appeal to a wide range of approaches – even though these approaches may explain and justify the norms differently.

To incorporate GR into our logical framework, we need to add requirements to be *conscientious* and *impartial*. Our belief logic already has part of the conscientiousness requirement. We already can prove the imperative analogue of the first step of our GR argument – "Don't *act* to do A to X without *believing* that it's all right for you to do A to X":[1]

$$
\begin{array}{rl}
 & [\ \therefore \sim(\underline{u}{:}A\underline{u}x \bullet \sim\underline{u}{:}RA\underline{u}x) \\
*\ 1 & \ulcorner\ \text{asm: } (\underline{u}{:}A\underline{u}x \bullet \sim\underline{u}{:}RA\underline{u}x) \\
2 & \mid\ \therefore \underline{u}{:}A\underline{u}x \quad \{\text{from } 1\} \\
*\ 3 & \mid\ \therefore \sim\underline{u}{:}RA\underline{u}x \quad \{\text{from } 1\} \\
*\ 4 & \mid\ u \therefore \sim RA\underline{u}x \quad \{\text{from } 3\} \\
5 & \mid\ u \therefore O\sim A\underline{u}x \quad \{\text{from } 4\} \\
6 & \mid\ u \therefore A\underline{u}x \quad \{\text{from } 2\} \\
7 & \llcorner\ u \therefore \sim A\underline{u}x \quad \{\text{from } 5\} \\
8 & \therefore \sim(\underline{u}{:}A\underline{u}x \bullet \sim\underline{u}{:}RA\underline{u}x) \quad \{\text{from 1; 6 contradicts 7}\}
\end{array}
$$

However, we can't yet prove the imperative analogue of our GR argument's third step – which also deals with conscientiousness:

> Don't *believe* that if you were in the same situation then it would be all right for X to do A to you, without *being willing* that if you were in the same situation then X do A to you.

The hard part here is to symbolize "in the same situation." If we ignore this for the moment, then what we need is this: "Don't *believe* that it would be all right for X to do A to you without *being willing* that X do A to you." We'll interpret "being willing that A be done" as "accepting 'A *may* be done.'" The permissive "A may be done" here isn't another way to say "A is all right." Instead, it's a member of the imperative family, but weaker than "Do A," expressing only one's consent to the action. We'll symbolize "A *may* be done" as "$M\underline{\mathbf{A}}$." Then we can symbolize the imperative mentioned above as follows:

[1] See my "Acting commits one to ethical beliefs," *Analysis* 42 (1983), pp. 40–3.

$\sim(\underline{\mathbf{u}}{:}RA\underline{\mathbf{x}}u \bullet \sim\underline{\mathbf{u}}{:}MA\underline{\mathbf{x}}u)$

Don't combine *believing* "It would be all right for X to do A to me" with *not accepting* "X may do A to me."

0323 To prove this, we need a principle like "$\Box(RA \supset MA)$" – which says that a permissibility judgment entails the corresponding permissive. This is like the prescriptivity principle ("Hare's Law") discussed in §12.4, which says that an ought judgment entails the corresponding imperative: "$\Box(O\underline{\mathbf{A}} \supset \underline{\mathbf{A}})$."[1]

Our biggest task is to symbolize and prove the impartiality requirement and the imperative analogue of our GR argument's second step:

$\sim(\underline{\mathbf{u}}{:}RA\underline{\mathbf{u}}x \bullet \sim\underline{\mathbf{u}}{:}\ldots)$

Don't combine *believing* that it would all right for you to do A to X with *not believing* that it if you were in the same situation then it would be all right for X to do A to you.

We need to replace "…" with a formula that means "it would be all right for X to do A to you in *the same situation*." And we need an inference rule to reflect universalizability – which is one of the few principles on whose truth almost all moral philosophers agree.

The *universalizability principle* (U) says that whatever is right (wrong, good, bad, etc.) in one case would also be right (wrong, good, bad, etc.) in any exactly or relevantly similar case, regardless of the individuals involved. Here are three equivalent formulations for "all right" (similar forms work for "ought"):

Universalizability (U)

If it's all right for X to do A, then it would be all right for anyone else to do A in the same situation.

If act A is permissible, then there is some universal property (or conjunction of such properties) F, such that: (1) act A is F, and (2) in any actual or hypothetical case every act that is F is permissible.

$$(R\underline{\mathbf{A}} \supset (\exists F)(F\underline{\mathbf{A}} \bullet \blacksquare(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset R\underline{\mathbf{X}})))$$

[1] On "$\Box(R\underline{A} \supset M\underline{A})$," see my "How incomplete is prescriptivism?" *Mind* 93 (1984), pp. 103–7. "$\Box(R\underline{A} \supset M\underline{A})$" and "$\Box(O\underline{A} \supset \underline{A})$" affirm that violating conscientiousness is logically inconsistent. One who rejected this but still thought that violating conscientiousness was objectionable, could endorse the weaker "$(R\underline{A} \supset M\underline{A})$" and "$(O\underline{A} \supset \underline{A})$" – and this would suffice for the GR proof at the end of this chapter.

The second phrasing, which is more technically precise, uses the notion of a "universal property." A *universal property* is any non-evaluative property describable without proper names (like "Gensler" or "Chicago") or pointer terms (like "I" or "this"). Suppose that I'm tempted to steal Pat's new computer. This possible act has several properties; for example, it's:

- *wrong* (evaluative term),
- an act of stealing *Pat's* computer (proper name), and
- something *I* would be doing (pointer word).

0324 These aren't universal, since they use evaluative terms, proper names, or pointer words. But the act also has universal properties; for example, it is:

- an act of stealing a new computer from one's neighbor,
- an act whose agent has blue eyes, and
- an act that would greatly distress the computer's owner.

U says that the morality of an act depends on its universal properties (like those of the second group), properties expressible without evaluative terms, proper names, or pointer words. Two acts with the same universal properties must have the same moral status, regardless of the individuals involved.

Here's an important corollary of universalizability:

> U* If it's all right for you to do A to X, then it would be all right for X to do A to you in the exact same situation.
>
> If it's all right for you to do A to X, then, for some universal property F, F is the complete description of your-doing-A-to-X in universal terms, and, in any actual or hypothetical case, if X's-doing-A-to-you is F, then it would be all right for X to do A to you.
>
> $(RA\underline{u}x \supset (\exists F)(F{*}A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset RA\underline{x}u)))$

U* relates closely to the second step in our argument for GR.

14.5 GR logical machinery

Now we add symbols for formulating GR:

- letters for universal properties and for actions,
- "M" ("may") for permissives,
- "■" ("in every actual or hypothetical case") for hypothetical cases, and

- “*” for the complete description of an act in universal terms.

We also add inference rules. This section will get complicated; you may need to read it a couple of times to follow what’s happening.

First, we’ll use letters of two new sorts (both can be used in quantifiers):

- “F,” “G,” “H,” and these with primes stand for universal properties of actions (including conjunctions of such properties).
- “$\underline{\mathbf{X}}$,” “$\underline{\mathbf{Y}}$,” “$\underline{\mathbf{Z}}$,” and these with primes stand for actions.

These examples use letters for universal properties:

0325

$F\underline{\mathbf{A}}$ = Act A is F (e.g., act A is an act of stealing)

Act A has universal property F

$(F\underline{\mathbf{A}} \supset \sim R\underline{\mathbf{A}})$ = If act A is an act of stealing, then act A is wrong

$G\underline{\mathbf{A}}$ = Act A is an act of a blue-eyed philosophy teacher stealing a bicycle from an impoverished student

We translate “$F\underline{\mathbf{A}}$” as “Act A is F” (not as “Imperative ‘Do A’ is F”). This next example uses a universal-property quantifier:

$(F)(F\underline{\mathbf{A}} \equiv F\underline{\mathbf{B}})$ = Acts A and B have all the same universal properties

For every universal property F, act A has property F if and only if act B has property F

These examples use action quantifiers:

$(\exists\underline{\mathbf{X}})F\underline{\mathbf{X}}$ = Some act has universal property F

For some act X, X has universal property F

$(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset O\underline{\mathbf{X}})$ = Every act that is F ought to be done

For every act X, if act X is F, then act X ought to be done

$(\underline{\mathbf{X}})(\exists F)F\underline{\mathbf{X}}$ = Every act has some universal property

For every act X there’s some universal property F, such that act X is F

These new symbols require new formation rules:

1. The result of writing "F," "G," "H," or one of these with primes, and then an imperative wff is itself a descriptive wff.

2. The result of writing "(x" or "(∃," and then "F," "G," "H," "X," "Y," "Z," or one of these with primes, and then ")" is a quantifier.

Assume expanded versions of our quantifier rules for the new quantifiers. We have to substitute the right sort of thing for the quantified letter:

1. For *individual variables*: x, y, z, x', ..., substitute individual constants: a, b, c, d, ...
2. For *universal-property variables*: F, G, H, F', ..., substitute universal-property letters not bound to quantifiers: F, G, H, F', ...
3. For *action variables*: **X**, **Y**, **Z**, **X**', ..., substitute imperative wffs: A**a**, **B**, A**x**y, ...[1]

0326

When "M" is prefixed to an imperative wff, we'll translate it as "may":[2]

3. The result of prefixing an imperative wff with "M" is a wff.

M**A** = Act A may be done
MA**x**u = X may do A to you

u:MA**x**u
= You accept "X may do A to me"
You consent to X's doing A to you
You're willing that X do A to you

Permissives like "M**A**" are weaker members of the imperative family. They express our consent to the act, but not necessarily our positive desire that the act take place. We can consistently consent both to the act and to its omission – saying "You may do A and you may omit A." Here are further wffs:

~M~**A** = Act A may not be omitted

[1] The last case requires two technical provisos. Suppose that we drop a quantifier containing an action variable and substitute an imperative wff for the variable. Then we must be sure that (1) this imperative wff contains no free variable that also occurs in a quantifier in the derived wff, and (2) if we dropped an existential quantifier, this substituted imperative wff must be an underlined capital letter that isn't an action variable and that hasn't occurred before in the proof.

[2] Capital letters have various uses, depending on context. In "((M • Ma) ⊃ (Mbc • MA))," for example, "M" is used first for a statement, then for a property of an individual, then for a relation between individuals, and finally for "may." It's usually clearer to use different letters.

u:~M~A**x**u
= You accept “X may not omit doing A to me”
You demand that X do A to you

“M**A**” is weaker and “~M~**A**” is stronger than “**A**.”[1]

Inference rule G1 is the principle that “A is all right” entails “A may be done.” G1 holds regardless of what imperative wff replaces “**A**”:[2]

G1
R**A** → M**A**

Given this and the rules for “M,” “O,” and “R,” we also can prove the reverse entailment from “M**A**” to “R**A**.” Then either logically entails the other; so accepting one commits us to accepting the other. But the distinction between the two doesn’t vanish. “R**A**” is true or false; to accept “R**A**” is to believe that something is true. But “M**A**” isn’t true or false; to accept “M**A**” isn’t to believe something but to will something, to consent to the idea of something being done.

Some of our inference rules for “M” (and later “■”) involve new kinds of world. A *world prefix* is now any string of zero-or-more instances of letters from 0327 the set <W, D, H, P, a, b, c, …> – where <a, b, c, …> is the set of small letters. Here “P,” “PP,” “PPP,” and so on are “permission worlds,” much like deontic worlds. A permission world that depends on a given world W1 is a possible world that contains the indicative judgments of W1 and some set of imperatives prescribing actions jointly permitted by the permissives of W1.

Inference rules G2 to G4 (which won’t be used in our GR proof) govern permissions and are much like the deontic rules. G2 and G3 hold regardless of what pair of contradictory imperative wffs replaces “**A**” / “~**A**”:

[1] For more on permissives, see my *Formal Ethics* (London: Routledge, 1996), pp. 185–6, and my “How incomplete is prescriptivism?” *Mind* 93 (1984), pp. 103–7.

[2] Thinking that an act is *all right* commits one to *consenting* to the idea of it being done (*being willing* that it be done). We also could use words like “approve,” “accept,” “condone,” or “tolerate” – in one sense of these terms. The sense of “consent” that I have in mind refers to an inner attitude incompatible with inwardly objecting to (condemning, disapproving, forbidding, protesting, objecting to) the act. Consenting here is a minimal attitude and needn’t involve favoring or advocating or welcoming the act. It’s consistent to both consent to the idea of A being done and also consent to the idea of A not being done.

G2

~M<u>**A**</u> → P ∴ ~<u>**A**</u>,
use a blank or any string of P's

In G2, the world prefix of the derived line must be either the same as that of the earlier line or else the same except that it adds one or more P's at the end.

G3

M<u>**A**</u> → P ∴ <u>**A**</u>,
use a *new* string of P's

In G3, the world prefix of the derived line must be the same as that of the earlier line except that it adds a *new* string (a string not occurring in earlier lines) of one or more P's at the end. G4 mirrors the deontic indicative transfer rule; it holds regardless of what descriptive or deontic wff replaces "A":

G4

P ∴ A → A

In G4, the world prefixes in the derived and deriving lines must be identical except that one ends in one or more additional P's.

"■" is a modal operator somewhat like "☐":

4. The result of prefixing any wff with "■" is a wff.

"■" translates as "in every actual or hypothetical case" or "in every possible world having the same basic moral principles as those true in the actual world." Here's a wff using "■":

■(F<u>**A**</u> ⊃ O<u>**A**</u>)
= If act A is or were F, then act A ought to be done
In every actual or hypothetical case, if act A is F, then act A ought to be done

Suppose that, while act A may or may not have property F (e.g., it may or may not maximize pleasure), still, if it did, then it would be what ought to be done. We'll use "$\blacksquare(F\underline{\mathbf{A}} \supset O\underline{\mathbf{A}})$" for this idea. "$(F\underline{\mathbf{A}} \supset O\underline{\mathbf{A}})$" is too weak to express this 0328 (since this wff is trivially true if "$F\underline{\mathbf{A}}$" is false); "$\Box(F\underline{\mathbf{A}} \supset O\underline{\mathbf{A}})$" is too strong (because there's no such entailment). So we'll use "$\blacksquare$" to formulate claims about what would be right or wrong in hypothetical situations (such as imagined exactly reversed situations).

We can now symbolize the *universalizability principle*:

> U If act A is permissible, then there's some universal property (or conjunction of such properties) F, such that: (1) act A is F, and (2) in any actual or hypothetical case every act that is F is permissible.
>
> $$(R\underline{\mathbf{A}} \supset (\exists F)(F\underline{\mathbf{A}} \bullet \blacksquare(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset R\underline{\mathbf{X}})))$$

G5 and G6 are the "all right" and "ought" forms of the corresponding inference rules. These hold regardless of what imperative wff replaces "$\underline{\mathbf{A}}$," what universal-property variable replaces "F," and what action variable replaces "$\underline{\mathbf{X}}$":

> G5 & G6
>
> $$R\underline{\mathbf{A}} \quad \rightarrow \quad (\exists F)(F\underline{\mathbf{A}} \bullet \blacksquare(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset R\underline{\mathbf{X}}))$$
> $$O\underline{\mathbf{A}} \quad \rightarrow \quad (\exists F)(F\underline{\mathbf{A}} \bullet \blacksquare(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset O\underline{\mathbf{X}}))$$

In G5 and G6, the world prefix of the derived and deriving lines must be identical and must contain no "W." The proviso prevents us from being able to prove the controversial idea that violations of universalizability are self-contradictory.

The rules for "$\blacksquare$" resemble those for "$\Box$." Recall that our expanded world prefixes can use "H," "HH," and "HHH"; these represent *hypothetical situation worlds*, which are possible worlds having the same basic moral principles as those of the actual world (or whatever world the H-world depends on). G7 and G8 hold regardless of what pair of contradictory wffs replaces "A" / "∼A":

> G7
>
> $$\blacksquare A \quad \rightarrow \quad H \therefore A,$$
> use a blank or any string of H's

In G7, the world prefixes in the derived and deriving lines must either be the same or be the same except that one adds one or more H's at the end.

G8
~■A → H ∴ ~A,
use a *new* string of H's

In G8, the derived line's world prefix must be the same as that of the earlier line except that it adds a *new* string (a string not occurring in earlier lines) of one or more H's at the end. Rule G9 (which won't be used in our GR proof) says that 0329 "☐" and "■" are equivalent when prefixed to descriptive wffs; this holds regardless of what *descriptive* wff replaces "A":

G9

■A ↔ ☐A

Our final symbol is "*"; this is used with universal-property letters to represent the complete description of an action in universal terms. Here's the rule for constructing wffs with "*," with an example:

5. The result of writing "F," "G," "H," or these with primes, then "*,"and then an imperative wff is itself a descriptive wff.

F*<u>**A**</u> = F is the complete description of act A in universal terms
F is the description of act A in universal terms which includes all the universal properties of act A

"F*<u>**A**</u>" means the same as this longer wff:

(F<u>**A**</u> • (G)(G<u>**A**</u> ⊃ ☐(<u>**X**</u>)(F<u>**X**</u> ⊃ G<u>**X**</u>)))
Act A is F, and every universal property G that A has is included as part of F
Act A is F, and, for every universal property G that A has, it's logically necessary that every act that's F also is G

We adopt the corresponding inference rule G10, which lets us go back and forth between "F*<u>**A**</u>" and this longer wff. G10 holds regardless of what distinct universal-property letters replace "F" and "G," what imperative wff replaces "<u>**A**</u>," and what action variable replaces "<u>**X**</u>":

G10

$$F^*\underline{\mathbf{A}} \quad \leftrightarrow \quad (F\underline{\mathbf{A}} \bullet (G)(G\underline{\mathbf{A}} \supset \Box(\underline{\mathbf{X}})(F\underline{\mathbf{X}} \supset G\underline{\mathbf{X}})))$$

Rule G11, our final inference rule, says that every act has a complete description in universal terms (even though it may be too long to write down). G11 is an axiom; it lets us put wff "$(\underline{\mathbf{X}})(\exists F)F^*\underline{\mathbf{X}}$" on any line of a proof:

G11

$$(\underline{\mathbf{X}})(\exists F)F^*\underline{\mathbf{X}}$$

We'll use "*" in symbolizing "exactly similar situation." Let "$A\underline{\mathbf{m}}x$" represent the act of my attacking X and "F" be its complete description:

$F^*A\underline{m}x$ = My-attacking-X has *complete universal description* F

Let's flesh this out. Let "G," "G'," … be my universal properties; these include
0330 properties like being a logician. Let "H," "H'," … be X's universal properties; these might include being an impoverished student. Let "R," "R'," … be the relationships between X and me; these might include X's being my student. Now property F would look like this, which describes the *actual situation*:

$FA\underline{m}x$ = My-attacking-X is an act of someone who is G, G', … attacking someone who is H, H', … and related to me in ways R, R', ….

Now we imagine an *exactly similar situation* if we imagine the situation where X's-attacking-me has this same description F:

$FA\underline{x}m$ = X's-attacking-me is an act of someone who is G, G', … attacking someone who is H, H', … and related to X in ways R, R', ….

In this imagined exactly similar situation, X is in my exact place – and I am in X's exact place. All our universal properties and relationships are switched.

We can now symbolize the reversed-situation corollary of universalizability:

U*. If it's all right for you to do A to X, then it would be all right for X to do A to you in the exact same situation.

If it's all right for you to do A to X, then, for some universal property F, F is the complete description of your-doing-A-to-X in universal terms, and, in any actual or hypothetical case, if X's-doing-A-to-you is F, then it would be all right for X to do A to you.

$(\mathrm{RA}\underline{\mathbf{u}}\mathrm{x} \supset (\exists \mathrm{F})(\mathrm{F^*A}\underline{\mathbf{u}}\mathrm{x} \bullet \blacksquare(\mathrm{FA}\underline{\mathbf{x}}\mathrm{u} \supset \mathrm{RA}\underline{\mathbf{x}}\mathrm{u})))$

Also, and most importantly, we can symbolize the golden rule:

GR. Treat others only as you consent to being treated in the same situation.

Don't combine *acting* to do A to X with *being unwilling* that if you were in the same situation then A be done to you.

Don't combine (1) accepting "Do A to X" with (2) not accepting "For some universal property F, F is the complete description in universal terms of my-doing-A-to-X, and, in any actual or hypothetical situation, if X's-doing-A-to-me is F, then X may do A to me."

$\sim(\underline{\mathbf{u}}:\mathrm{A}\underline{\mathbf{u}}\mathrm{x} \bullet \sim\underline{\mathbf{u}}:(\exists \mathrm{F})(\mathrm{F^*A}\underline{\mathbf{u}}\mathrm{x} \bullet \blacksquare(\mathrm{FA}\underline{\mathbf{x}}\mathrm{u} \supset \mathrm{MA}\underline{\mathbf{x}}\mathrm{u})))$

Here are symbolizations of two related ideas (§14.2): 0331

Impartiality: Make similar evaluations about similar actions, regardless of the individuals involved.

Don't accept "Act A is permissible" without accepting "Any act exactly or relevantly similar to act A is permissible."

Don't accept "Act A is permissible" without accepting "For some universal property F, act A is F and, in any actual or hypothetical situation, any act that is F is permissible."

$\sim(\underline{\mathbf{u}}:\mathrm{R}\underline{\mathbf{A}} \bullet \sim\underline{\mathbf{u}}:(\exists \mathrm{F})(\mathrm{F}\underline{\mathbf{A}} \bullet \blacksquare(\underline{\mathbf{X}})(\mathrm{F}\underline{\mathbf{X}} \supset \mathrm{R}\underline{\mathbf{X}})))$

Formula of universal law: Act only as you're willing for anyone to act in the same situation – regardless of imagined variations of time or person.[1]

Don't combine *acting* to do A with *not being willing* that any similar action be done in the same situation.

Don't combine (1) accepting "Do A" with (2) not accepting "For some universal property F, F is the complete description in universal terms of my doing A, and, in any actual or hypothetical situation, any act that is F may be done."

[1] My "formula of universal law" resembles Immanuel Kant's principle. His wording went, "Act only on that maxim through which you can at the same time will that it should be a universal law." I'm not claiming that Kant explicitly intended his principle in exactly my sense.

$\sim(\underline{u}{:}A\underline{u} \bullet \sim\underline{u}{:}(\exists F)(F^*A\underline{u} \bullet \blacksquare(\underline{X})(F\underline{X} \supset M\underline{X})))$

This “formula of universal law” is a generalized GR. It applies, for example, to multi-party cases or to cases where my present action can harm my future self.

14.6 The symbolic GR proof

Before we do our GR proof, let’s review the larger picture.

We began this chapter by sketching various dimensions of ethical rationality. Then we narrowed our focus, first to consistency, and then to a single consistency principle – the golden rule. We had to formulate GR carefully to avoid absurd implications. We defended this wording:

> *Golden rule*
>
> Treat others only as you consent to being treated in the same situation.
>
> GR forbids this combination:
>
> - I do A to another.
> - I’m unwilling that if I were in the same situation then A be done to me.

We sketched an intuitive GR proof, using the example of stealing Detra’s bicycle. 0332 Then we noted that incorporating GR and its proof into our logical framework requires adding impartiality and strengthening conscientiousness. And so now we’re ready to give a formal proof of the golden rule.

Our proof goes as follows (where justifications that use our new inference rules are in **bold type**):

[$\therefore \sim(\underline{u}{:}A\underline{u}x \bullet \sim\underline{u}{:}(\exists F)(F^*A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u)))$

1 asm: $(\underline{u}{:}A\underline{u}x \bullet \sim\underline{u}{:}(\exists F)(F^*A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u)))$

2 $\therefore \underline{u}{:}A\underline{u}x$ {from 1}

3 $\therefore \sim\underline{u}{:}(\exists F)(F^*A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u))$ {from 1}

4 $u \therefore \sim(\exists F)(F^*A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u))$ {from 3}

5 $u \therefore A\underline{u}x$ {from 2}

6 u asm: $\sim RA\underline{u}x$ {we need to derive “$RA\underline{u}x$”}

7 $u \therefore O\sim A\underline{u}x$ {from 6}

8 $u \therefore \sim A\underline{u}x$ {from 7}

9 $u \therefore RA\underline{u}x$ {from 6; 5 contradicts 8}

10 $u \therefore (\exists F)(FA\underline{u}x \bullet \blacksquare(\underline{X})(F\underline{X} \supset R\underline{X}))$ **{from 9 by G5}**

11 $u \therefore (GA\underline{u}x \bullet \blacksquare(\underline{X})(G\underline{X} \supset R\underline{X}))$ {from 10}

12 $u \therefore GA\underline{u}x$ {from 11}

13 $u \therefore \blacksquare(\underline{X})(G\underline{X} \supset R\underline{X})$ {from 11}

14 $u \therefore (\underline{X})(\exists F)F^*\underline{X}$ **{rule G11}**

15 $u \therefore (\exists F)F{*}A\underline{u}x$ {from 14}
16 $u \therefore H{*}A\underline{u}x$ {from 15}
17 $u \therefore (HA\underline{u}x \bullet (F)(FA\underline{u}x \supset \Box(\underline{X})(H\underline{X} \supset F\underline{X})))$ **{from 16 by G10}**
18 $u \therefore HA\underline{u}x$ {from 17}
19 $u \therefore (F)(FA\underline{u}x \supset \Box(\underline{X})(H\underline{X} \supset F\underline{X}))$ {from 17}
20 $u \therefore (GA\underline{u}x \supset \Box(\underline{X})(H\underline{X} \supset G\underline{X}))$ {from 19}
21 $u \therefore \Box(\underline{X})(H\underline{X} \supset G\underline{X})$ {from 12 and 20}
22 $u \therefore (F){\sim}(F{*}A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u))$ {from 4}
23 $u \therefore {\sim}(H{*}A\underline{u}x \bullet \blacksquare(HA\underline{x}u \supset MA\underline{x}u))$ {from 22}
24 $u \therefore {\sim}\blacksquare(HA\underline{x}u \supset MA\underline{x}u)$ {from 16 and 23}
25 $uH \therefore {\sim}(HA\underline{x}u \supset MA\underline{x}u)$ **{from 24 by G8}**
26 $uH \therefore HA\underline{x}u$ {from 25}
27 $uH \therefore {\sim}MA\underline{x}u$ {from 25}
28 $uH \therefore (\underline{X})(H\underline{X} \supset G\underline{X})$ {from 21}
29 $uH \therefore (HA\underline{x}u \supset GA\underline{x}u)$ {from 28}
30 $uH \therefore GA\underline{x}u$ {from 26 and 29}
31 $uH \therefore (\underline{X})(G\underline{X} \supset R\underline{X})$ **{from 13 by G7}**
32 $uH \therefore (GA\underline{x}u \supset RA\underline{x}u)$ {from 31}
33 $uH \therefore RA\underline{x}u$ {from 30 and 32}
34 $uH \therefore MA\underline{x}u$ **{from 33 by G1}**
35 $\therefore {\sim}(\underline{u}{:}A\underline{u}x \bullet {\sim}\underline{u}{:}(\exists F)(F{*}A\underline{u}x \bullet \blacksquare(FA\underline{x}u \supset MA\underline{x}u)))$ {from 1; 27 contradicts 34}

While this is a difficult proof, you should be able to follow the individual lines and see that everything follows correctly.

Our proof begins as usual; we assume the opposite of what we want to prove and then try to derive a contradiction. Soon we get lines 4 and 5 (where 5 is 0333 addressed to yourself):

4 X may not do A to me in an exactly similar situation.
5 Do A to X.

Using line 4, we get these key lines:

16 Let H be the complete description of my doing A to X.
26 In our imagined situation, X's-doing-A-to-me is H.
27 In our imagined situation, X may not do A to me.

We use line 5 to get "It's all right for me to do A to X":

5 Do A to X. (This is addressed to yourself.)
6 Assume that it's not all right for me to do A to X.
7 ∴ I ought not to do A to X.
8 ∴ Don't do A to X. (This is addressed to yourself.)
9 ∴ It's all right for me to do to X. {8 contradicts 5}

Then we use universalizability on "It's all right for me to do A to X" to get "Any act relevantly or exactly similar to my-doing-A-to-X would be all right." We specify that G is the morally relevant complex of properties here; so:

12 My-doing-A-to-X has property G.
13 Any act that has property G would be all right.

We get a contradiction between lines 27 and 34:

16 H is the complete description of my doing A to X. {above}
12 My-doing-A to-X has property G. {above}
21 ∴ G is part of H – so every act that is H is G. {from 16 & 12}
26 In our imagined situation, X's-doing-A-to-me is H. {above}
30 ∴ In our imagined situation, X's-doing-A-to-me is G. {from 21 & 26}
13 Any act that has property G would be all right. {above}
33 ∴ In our imagined situation, X's-doing-A-to-me is all right. {from 30 & 13}
34 ∴ In our imagined situation, X may do A to me. {from 33}

Thus ends our proof of the golden rule:[1]

Always treat others as you want to be treated; that is the summary of the Law and the Prophets. (Mt 7:12)

[1] For a challenging exercise, prove the impartiality and universal law formulas, as formulated at the end of the previous section. Answers are in the back of the book.

15 Metalogic

Metalogic studies logical systems. It focuses on proving things about the systems themselves, not on testing concrete arguments. This chapter gives a brief introduction to metalogic.

15.1 Metalogical questions

Metalogic is the study of logical systems and tries to prove things about them. Recall our first two rules in §6.1 for forming propositional wffs:

1. Any capital letter is a wff.
2. The result of prefixing any wff with "∼" is a wff.

It follows from these that there's no longest wff – since, if there were a longest wff, then we could make a longer one by adding another "∼." This simple proof is about a logical system, so it's part of *metalogic*.

Consider our system of propositional logic. Metalogic asks questions like: Do we need all five symbols ("∼," "•," "∨," "⊃," and "≡")? Could we define some symbols in terms of others? Did we set up our proof system right? Are any of the inference rules defective? Can we prove self-contradictions or invalid arguments? Do we have enough inference rules to prove all valid propositional arguments? What other approaches could systematize propositional logic?

15.2 Symbols

We don't need all five propositional symbols ("∼," "•," "∨," "⊃," and "≡"). We could symbolize and test the same arguments if we had just "∼" and "•"; then, instead of writing "(P ∨ Q)," we could write "∼(∼P • ∼Q)":

(P ∨ Q) = ∼(∼P • ∼Q)
At least one is true = Not both are false

These are equivalent (true or false under the same conditions); truth tables can 0335 show this. Similarly, we can express “⊃” and “≡” using “∼” and “•”:

(P ⊃ Q) = ∼(P • ∼Q)
If P then Q = We don’t have P true and Q false

(P ≡ Q) = (∼(P • ∼Q) • ∼(Q • ∼P))
P if and only if Q = We don’t have P true and Q false, and we don’t have Q true and P false

Or we might translate the other symbols into “∼” and “∨”:

(P • Q) = ∼(∼P ∨ ∼Q)
(P ⊃ Q) = (∼P ∨ Q)
(P ≡ Q) = (∼(P ∨ Q) ∨ ∼(∼P ∨ ∼Q))

Or we might use just “∼” and “⊃”:

(P • Q) = ∼(P ⊃ ∼Q)
(P ∨ Q) = (∼P ⊃ Q)
(P ≡ Q) = ∼((P ⊃ Q) ⊃ ∼(Q ⊃ P))

It’s possible to get by using just “|” for NAND; “(P | Q)” means “not both P and Q.” We can define “∼P” as “(P | P)” and “(P • Q)” as “((P | Q) | (P | Q)).”

Systems with only one or two symbols are more elegantly simple but harder to use. But logicians are sometimes more interested in proving results about a system than in using it to test arguments; and it may be easier to prove these results if we use fewer symbols.

Another approach uses all five symbols but divides them into undefined (primitive) symbols and defined ones. We could take “∼” and either “•” or “∨” or “⊃” as undefined, and then define the others using these. We’d then view the defined symbols as abbreviations; whenever we liked, we could eliminate them and use only undefined symbols.

How do we know that our five symbols suffice to formulate wffs for every possible truth table? Suppose we have a truth table for two letters that comes out as below and we want to replace “??” with a wff that gives this table:

A B	??
0 0	0
0 1	1
1 0	1
1 1	0

How do we know that some wff gives this truth table? To construct a wff with this truth table, we can put an OR between the true cases (rows 2 and 3): A-is-false-and-B-is-true (row 2) OR A-is-true-and-B-is-false (row 3):

((∼A • B) ∨ (A • ∼B)) 0336

So we can, using just NOT, AND, and OR, mechanically construct a wff that expresses any specific truth table. (If the formula is always false, use a wff like "(A • ∼A)," which is always false.)

There are further options about notation. While we use capital letters for statements, some logicians use small letters (often just "p," "q," "r," and "s") or Greek letters. Some use "-" or "¬" for negation, "&" or "∧" for conjunction, "→" for conditional, or "↔" for equivalence. Various conventions are used for dropping parentheses. It's easy to adapt to these differences.

Polish notation avoids parentheses and has shorter formulas. "K," "A," "C," and "E" go in place of the left-hand parentheses for "•," "∨," "⊃," and "≡"; and "N" is used for "∼." Here are four examples:

∼(P • Q)	=	NKpq
(∼P • Q)	=	KNpq
((P • Q) ⊃ R)	=	CKpqr
(P • (Q ⊃ R))	=	KpCqr

Some people can actually understand these formulas.

15.3 Soundness

The most important metalogical questions are about whether a proof system is **sound** (won't prove bad things – so every argument provable in the system is valid) and **complete** (can prove every good thing – so every valid argument expressible in the system is provable in the system).

Could the following happen? A student named Logicus found a flaw in our proof system. Logicus did a formal proof of a propositional argument and then showed by a truth table that the argument was invalid; so some arguments provable using our proof system are invalid. People have found such flaws in proof systems. How do we know that our system is free from such flaws? How can we prove soundness?

Soundness: Any propositional argument for which we can give a formal proof is valid (on the truth-table test).

To show this, we could first show that all the propositional inference rules are *truth preserving* (when applied to true wffs, they yield only further true wffs). We have 13 inference rules: 6 S-rules, 6 I-rules, and RAA. It's easy (but tedious) to use the truth-table method of §6.6 to show that S- and I-rules are truth preserving. All these rules pass the test (as you could check for yourself); when applied to true wffs, they yield only further true wffs.

RAA is more difficult to check. First we show that the *first* use of RAA in a proof is truth preserving. Suppose all previous not-blocked-off lines in a proof are true, and we use RAA to derive a further line; we have to show that this further line is true: 0337

. . . .
    asm: ∼A
    . . . .
    ∴ **B**
    ∴ **∼B**
∴ A

From previous true lines plus assumption "∼A," we derive contradictory wffs "B" and "∼B" using S- and I-rules. We just saw that S- and I-rules are truth preserving. So if the lines used to derive "B" and "∼B" were all true, then both "B" and "∼B" would have to be true, which is impossible. Hence the lines used to derive them can't all be true. So if the lines before the assumption are all true, then assumption "∼A" has to be false. So its opposite ("A") has to be true. So the first use of RAA in a proof is truth preserving.

We can similarly show that if the first use of RAA in a proof is truth preserving, then the second is too. And we can show that if the first n uses of RAA are truth preserving, then the n + 1 use is too. Then we can apply the principle of *mathematical induction*: "Suppose that something holds in the first case, and that, if it holds in the first n cases, then it holds in the n + 1 case; then it holds in all cases." From this, it follows that *all* uses of RAA are truth preserving.

Now suppose an argument is provable in our propositional system. Then there's some proof that derives the conclusion from the premises using truth-preserving rules. So if the premises are true, then the conclusion also must be true – and so the argument is valid. So if an argument is provable in our propositional system, then it's valid. This establishes soundness.

Isn't this reasoning circular? Aren't we just assuming principles of propositional inference (like *modus ponens*) as we defend our propositional system? Of course we are. Nothing can be proved without assuming logical rules. We aren't attempting the impossible task of proving things about a logical system without assuming any logical rules. Instead, we're doing something more modest. We're trying to show, relying on ordinary reasoning, that we didn't make errors in setting up our system.

The consistency of our system is an easy corollary of its soundness. Let's say that a wff is a **theorem** if it's provable from zero premises. "(P ∨ ~P)" is a theorem; we can prove it by assuming its opposite and deriving a contradiction:

[∴ (P ∨ ~P)   **Valid**
* 1 ┌ asm: ~(P ∨ ~P)
2 │ ∴ ~P   {from 1}
3 └ ∴ P   {from 1}
4 ∴ (P ∨ ~P)   {from 1; 2 contradicts 3}

By our soundness result, since "∴ (P ∨ ~P)" is provable it must be valid on the truth-table test. So then it must be impossible for "(P ∨ ~P)" to be false. So then "(P ∨ ~P)" must have an all-1 truth table. And the more general result follows, 0338 that all theorems of our system must have all-1 truth tables.

A proof system is **consistent** provided that no two contradictory formulas are both theorems. We showed that all theorems of our system have all-1 truth tables. But no two contradictory formulas both have all-1 truth tables (since if a formula has all 1's then its contradictory has all 0's). So no two contradictory formulas are both theorems. So our propositional system is consistent.

15.4 Completeness

Our soundness proof shows that our propositional system won't prove invalid arguments. You probably didn't doubt this. But you may have had doubts about whether our system is strong enough to prove all valid propositional arguments. After all, the single-assumption method of doing proofs wasn't strong enough; Section 7.3 uncovered valid arguments that require multiple assumptions. How do we know that our expanded method is enough? Maybe Logicus will find a further propositional argument that's valid but not provable; then we'd have to strengthen our system still further. To calm these doubts, we'll show that our propositional system is complete:

Completeness: Every valid propositional argument is provable.

Our completeness proof will show that if we correctly apply the proof strategy of §7.3 to a valid propositional argument then we get a proof. Our strategy has five steps: 1-START, 2-S&I, 3-RAA, 4-ASSUME, and 5-REFUTE. Assume that we correctly apply this strategy to a propositional argument. Then:

We’ll end in the RAA step with all assumptions blocked off, or end in the REFUTE step, or keep going endlessly.
If we end in the RAA step with all assumptions blocked off, then we’ll get a proof.
If we end in the REFUTE step, then the argument is invalid.
We won’t keep going endlessly.
∴ If the argument is valid, then we’ll get a proof.

((A ∨ F) ∨ E)
(A ⊃ P)
(F ⊃ ∼V)
∼E
∴ (V ⊃ P)

Premise 1 is true because our proof strategy has only two stopping points; so we’ll stop at one or the other or we won’t stop. Premise 2 is true because our proof strategy (especially the S&I and RAA steps) mirrors the §7.1 definition of “proof.” Now we have to argue for premises 3 and 4.

Premise 3 says “If we end in the REFUTE step, then the argument is invalid.” This is true because, when we reach the REFUTE step, all the complex wffs are dissolved into smaller parts and eventually into simple wffs, the larger forms are true if the smaller parts are true, and the simple wffs we end up with are consistent and thus give truth conditions making all the other wffs true – thus 0339 making the premises of the original argument true while its conclusion is false – thus showing that the original argument is invalid.

Here’s a chart (where α and β represent wffs) about how complex wff forms would dissolve into simpler wff forms:

∼∼α dissolves into α [S-rule]
(α • β) dissolves into α and β [S-rule]
∼(α • β) dissolves into ∼α or ∼β [I-rule or assumption]
(α ∨ β) dissolves into α or β [I-rule or assumption]
∼(α ∨ β) dissolves into ∼α and ∼β [S-rule]
(α ⊃ β) dissolves into ∼α or β [I-rule or assumption]
∼(α ⊃ β) dissolves into α and ∼β [S-rule]
(α ≡ β) dissolves into (α ⊃ β) and (β ⊃ α) [S-rule]
∼(α ≡ β) dissolves into (α ∨ β) and ∼(α • β) [S-rule]

The original formula is true if the parts it dissolves into are true.

The chart covers the nine complex wff forms possible in our system and the smaller parts that these complex wff forms will have dissolved into when we reach the REFUTE step. Forms that dissolve using an S-rule always dissolve into the same smaller parts. Other forms can dissolve in two ways. Consider “∼(A • B).” We might be able to use an I-rule on this to derive “∼A” or “∼B.” If not, then we can break "$\sim(A \bullet B)$" by assuming one part or its negation, which will (immediately or after using an I-rule) give us "$\sim A$" or "$\sim B$." So when we reach the REFUTE step, all not-blocked-off complex wffs will be *starred* or *broken*,[1] and thus dissolved into the parts given above.

Each complex wff is true if the parts it dissolves into are true. We can check this by checking the nine cases in the box. So $\sim\sim\alpha$ dissolves into α, and is true if α is true. $(\alpha \bullet \beta)$ dissolves into α and β, and is true if both of these are true. Similarly, $\sim(\alpha \bullet \beta)$ goes into $\sim\alpha$ or $\sim\beta$, and is true if either of these is true.

Our *refutation* is the set of all the simple not-blocked-off wffs and is consistent (or else we'd have applied RAA). This refutation gives truth conditions making all the other not-blocked-off wffs true too (since these other wffs dissolved into the simple parts that make up the refutation). So our refutation gives truth conditions making *all* the not-blocked-off lines true. But these lines include the premises and the denial of the conclusion (of the original argument). So our refutation gives truth conditions making the premises and the denial of the conclusion all true. So the argument is *invalid*. So if we correctly apply our strategy to a propositional argument and end in the REFUTE step, then the argument is invalid. This establishes premise 3.

Now we argue for premise 4: "We won't keep going endlessly." This is a concern, since the proof strategy for some systems can go into an endless loop (§9.5). 0340 That won't happen in propositional logic, since here the complexity of the wffs that are neither starred nor blocked off nor broken keeps decreasing as we go on, and eventually, if we don't get a proof, goes to zero, at which point we get a refutation. For the tedious details, study the next paragraph.

Let the *complexity level of a wff* be the number of instances of "$\bullet$," "$\vee$," "$\supset$," and "$\sim\sim$" (double negation) that the wff would have if every wff in it of the form "$(\alpha \equiv \beta)$" were replaced with "$((\alpha \supset \beta) \bullet (\beta \supset \alpha))$." So *simple wffs* "$A$" and "$\sim A$" have complexity 0, "$(P \bullet Q)$," "$\sim(P \vee Q)$," "$\sim(\sim P \supset \sim Q)$," and "$\sim\sim P$" have complexity 1, "$((P \bullet Q) \supset R)$" has complexity 2, and "$(P \equiv (Q \vee R))$" has complexity 5. The *complexity level of a stage of a proof* is the sum of the complexity levels of the lines to that point that aren't either starred or blocked off or broken. When we START by assuming the conclusion's opposite, the argument has a certain complexity level; the sample problem at the start of Chapter 7 has complexity 3. Each S&I step (for example, going from "$(P \supset Q)$" and "P" to "Q" – or from "$(P \equiv Q)$" to "$(P \supset Q)$" and "$(Q \supset P)$") decreases the complexity level by at least one.[2] Any further ASSUME will immediately or in the next step (through an application of an I-rule) reduce

[1] A wff is *broken* if we already have one side or its negation but not what we need to conclude anything new (§7.3).

[2] One rare occasions, an S&I step can reduce the complexity level by more than one. Suppose that we have "$(A \bullet B)$" and "$(B \bullet A)$" and simplify one of them into "A" and "B." The conjunction we simplify is *starred* and the other one is *broken*, so the complexity level is reduced by two.

the complexity level by at least one.[1] RAA is trickier. If we apply RAA on the initial assumption, then the proof is done and there's no endless loop. If we apply RAA on a non-initial assumption, then the complexity level may temporarily increase (due to our having to erase multiple stars); but the overall effect is to decrease the complexity from what it was before we made the non-initial assumption in question.[2] So the complexity level keeps decreasing. Since the proof starts with a finite complexity level which keeps going down, then, if we don't get a proof, then we'll eventually end with a complexity level of 0 – which (if we can derive nothing further) will move us to the REFUTE step which ends the strategy. So we won't get an endless loop.

So if we correctly apply our strategy to a propositional argument and the argument is valid, then we'll get a proof. This establishes completeness. So we've proved both *soundness* (every provable propositional argument is valid) and *completeness* (every valid propositional argument is provable) for our system. 0341

15.5 An axiomatic system

Our propositional system is an *inferential system*, since it uses mostly **inference rules** (these let us derive formulas from earlier formulas). It's also possible to systematize propositional logic as an *axiomatic system*, which uses mostly **axioms** (formulas that can be put on any line, regardless of earlier lines). Both approaches can be equally powerful: anything provable with one is provable with the other. Axiomatic systems have a simpler structure while inferential systems are easier to use. Symbolic logic's pioneers used axiomatic systems.

I'll now sketch a version of an axiomatic system from *Principia Mathematica*.[3] We'll use our earlier definitions of "wff," "premise," and "derived line." But now a *proof* is a vertical sequence of zero or more premises followed by one or more derived lines, where each derived line is an axiom or follows from earlier lines by the inference rule or the substitution of definitional equivalents. There are four axioms; these axioms, and the inference rule and

[1] Suppose we need to break "(A ⊃ B)" and so we assume "A"; then we can conclude "B" and star "(A ⊃ B)," which will reduce the complexity by one. Suppose that instead we assume "∼A"; then "(A ⊃ B)" is broken, which immediately reduces the complexity by one.

[2] Suppose we make an additional assumption to break a complex wff. For example, we assume "A" to break "(A ⊃ B)." If this assumption dies, we conclude "∼A" and then "(A ⊃ B)" is broken (which reduces the complexity level). If instead we assumed "∼A," then when this assumption dies then we derive "A"; we then can use this with "(A ⊃ B)" to get "B" – and then star "(A ⊃ B)" (which reduces the complexity level). So when an additional assumption dies, then the complexity level is decreased from what it was before we made the assumption.

[3] Bertrand Russell and Alfred North Whitehead (Cambridge: Cambridge University Press, 1910).

definitions, hold regardless of which wffs uniformly replace "A," "B," and "C":

Axiom 1. ((A ∨ A) ⊃ A)
Axiom 2. (A ⊃ (A ∨ B))
Axiom 3. ((A ∨ B) ⊃ (B ∨ A))
Axiom 4. ((A ⊃ B) ⊃ ((C ∨ A) ⊃ (C ∨ B)))

The system has one inference rule (*modus ponens*): "(A ⊃ B), A → B." It takes "∨" and "∼" as undefined; it defines "⊃," "•," and "≡" as follows:

Definition 1. (A ⊃ B) = (∼A ∨ B)
Definition 2. (A • B) = ∼(∼A ∨ ∼B)
Definition 3. (A ≡ B) = ((A ⊃ B) • (B ⊃ A))

The inferential proof of "(P ∨ ∼P)" in our system is trivially simple (§15.3). The axiomatic proof is difficult:

1 ∴ (((P ∨ P) ⊃ P) ⊃ ((∼P ∨ (P ∨ P)) ⊃ (∼P ∨ P))) {from axiom 4, substituting "(P ∨ P)" for "A," "P" for "B," and "∼P" for "C"}
2 ∴ ((P ∨ P) ⊃ P) {from axiom 1, substituting "P" for "A"}
3 ∴ ((∼P ∨ (P ∨ P)) ⊃ (∼P ∨ P)) {from 1 and 2}
4 ∴ (P ⊃ (P ∨ P)) {from axiom 2, substituting "P" for "A" and "P" for "B"}
5 ∴ (∼P ∨ (P ∨ P)) {from 4, substituting things equivalent by definition 1}
6 ∴ (∼P ∨ P) {from 3 and 5}
7 ∴ ((∼P ∨ P) ⊃ (P ∨ ∼P)) {from axiom 3, substituting "∼P" for "A" and "P" for "B"}
8 ∴ (P ∨ ∼P) {from 6 and 7} 0342

Since there's no automatic strategy, creating such proofs requires guesswork and intuition. And we might work for hours trying to prove an argument that's actually invalid. Axiomatic systems tend to be painful to use.

15.6 Gödel's theorem

Now we'll consider metalogic's most surprising discovery: Gödel's theorem.

Let's define a **formal system** (or *calculus*) to be an artificial language with notational grammar rules and notational rules for determining validity, where these rules can be applied in a mechanical way to give a definite result about wffhood and validity in a finite amount of time. Many formal systems are inferential (our approach) or axiomatic.

Propositional logic can be put into a sound and complete formal system. Our inferential system does the job – as does the axiomatic system we just considered. In either, an argument is valid if and only if it's provable.

You might think that arithmetic could similarly be put into a sound and complete system. If we succeeded, we'd have an inferential or axiomatic system that could prove any truth of arithmetic but no falsehood. Then a statement of arithmetic would be true if and only if it's provable in the system.

But this is impossible. **Gödel's theorem** shows that we can't systematize arithmetic in this way. For any attempted formalization, one of two bad things will happen: some true statements of arithmetic won't be provable (making the system incomplete), or some false statements of arithmetic will be provable (making the system unsound). Gödel's theorem shows that any formal system of arithmetic will be incomplete or unsound.

You may find Gödel's theorem hard to believe. Arithmetic seems to be an area where everything can be proved one way or the other. But Kurt Gödel in 1931 showed that this was wrong. The reasoning behind his theorem is difficult; here I'll just try to give a glimpse of what it's about.[1]

What is this "arithmetic" that we can't systematize? "Arithmetic" here is roughly like high-school algebra, but limited to positive whole numbers. It includes truths like these three:

$2 + 2 = 4$
If $x + y = z$, then $y + x = z$.
If $xy = 18$ and $x = 2y$, then $x = 6$ and $y = 3$.

More precisely, *arithmetic* is the set of truths and falsehoods that can be expressed using symbols for the vocabulary items in these boxes: 0343

Mathematical vocabulary
positive numbers: 1, 2, … plus, times, to the power of parentheses, equals
Logical vocabulary
not, and, or, if-then variables (x, y, …), all, some parentheses, equals

Gödel's theorem claims that no formal system with symbols for all the items in these two boxes can be both sound and complete.

[1] My little *Gödel's Theorem Simplified* (Langham, Md.: University Press of America, 1984) tries to explain the theorem. Refer to this book for further information.

The notions in our mathematical box can be reduced to a sound and complete formal system; we'll call it the "number calculus." And the notions in our logical box can be reduced to a sound and complete formal system: our quantificational system. But combining these two systems produces a monster that can't be put into a sound and complete formal system.

We'll now construct a *number calculus* (NC) that uses seven symbols:

/ + • ^ ( ) =

"/" means "one" ("1"). We'll write 2 as "//" ("one one"), 3 as "///" ("one one one"), and so on. "+" is for "plus," "•" for "times," and "^" for "to the power of." Our seven symbols express all the notions in our mathematical box.

Meaningful sequences of NC symbols are *numerals*, *terms*, and *wffs*:

1. Any string consisting of one or more instances of "/" is a numeral.
2. Every numeral is a term.
3. The result of joining any two terms by "+," "•," or "^" and enclosing the result in parentheses is a term.
4. The result of joining any two terms by "=a" is a wff.

Here are examples (with equivalents):

- 2, 4 (*numerals*): // ////
- 2, 2 • 2, $(1 + 1)^2$ (*terms*): // (// • //) ((/ + /) ^ //)
- 4 = 4, 2 + 2 = 4 (*wffs*): //// = //// (// + //) = ////

Our NC can prove the true wffs. NC uses one axiom and six inference rules; here's our axiom (in which any numeral can replace "a"):

Axiom: a = a

Any instance of this (any self-identity using the same numeral on both sides) is an axiom: "/=/," "//=//," "///=///,"and so on.

Our inference rules let us substitute one string of symbols for another. We'll use "↔" to say that we can substitute the symbols on either side for those on the other side. We have two rules for "plus" (where "a" and "b" in our inference 0344 rules stand for any numerals):

R1. (a+/) ↔ a/
R2. (a+/b) ↔ (a/+b)

R1 lets us interchange "(///+/)"and "////." R2 lets us interchange "(//+//)"and "(///+/)" – moving the "+" one "/" to the right. We'll see R3 to R6 in a moment.

An *NC proof* is a vertical sequence of wffs, each of which is either an axiom or else follows from earlier members by one of the inference rules R1 to R6. A *theorem* is any wff of a proof.

Using our axiom and inference rules R1 and R2, we can prove any true wff of NC that doesn't use "•" or "^." Here's a proof of "(//+//)=////" ["2 + 2 = 4"]:

1. ////=////   {from the axiom}
2. (///+/)=////   {from 1 using R1}
3. (//+//)=////   {from 2 using R2}

We start with a self-identity. We get line 2 by substituting "(///+/)" for "////" (as permitted by rule R1). We get line 3 by further substituting "(//+//)" for "(///+/)" (as permitted by rule R2). So "(//+//)=////" is a theorem.

Here are our rules for "times" and "to the power of":

R3. (a • /) ↔ a
R4. (a • /b) ↔ ((a • b) + a)
R5. (a ^ /) ↔ a
R6. (a ^ /b) ↔ ((a ^ b) • a)

Our NC is sound and complete; any wff of NC is true if and only if it's provable in NC. This is easy to show, but we won't do the proof here.

Suppose we take our number calculus, add the symbols and inference rules of our quantificational logic, add a few more axioms and inference rules, and call the result the "arithmetic calculus" (AC). We could then symbolize any statement of arithmetic in AC. So we could symbolize these:

If x + y = z, then y + x = z.
((x+y)=z ⊃ (y+x)=z)

If xy = 8 and x = 2y, then x = 4 and y = 2.
(((x•y)=//////// • x=(//•y)) ⊃ (x=//// • y=//))

x is even.
For some number y, x = 2 times y.
(∃y)x=(// • y)

x is prime.
For every number y and z, if x = y times z, then y = 1 or z = 1.
$(y)(z)(x=(y \bullet z) \supset (y=/ \vee z=/))$ 0345

Here’s Goldbach’s conjecture (which is still neither proved nor disproved):

Every even number is the sum of two primes.

$(x)((\exists y)x=(2 \bullet y) \supset (\exists x')(\exists x'')(x=(x'+x'') \bullet ((y)(z)(x'=(y \bullet z) \supset (y=/ \vee z=/)) \bullet$
$(y)(z)(x''=(y \bullet z) \supset (y=/ \vee z=/)))))$

Gödel’s theorem shows that any such arithmetic calculus has a fatal flaw: either it *can’t* prove some arithmetic truths or it *can* prove some arithmetic falsehoods. This flaw comes not from an accidental defect in our choice of axioms and rules, but from the fact that any such system can encode messages about itself.

To show how this works, it’s helpful to use a version of AC with minimal vocabulary. The version that we’ve sketched so far uses these symbols:

/   +   •   ^   (   )   =   ~   ∨   ⊃   ∃   x, y, z, x’, ...

We’ll now economize. Instead of writing “^” (“to the power of”), we’ll write “••.” We’ll drop “∨” and “⊃,” and express the same ideas using “~” and “•” (§15.2). We’ll use “n,” “nn,” “nnn,” “nnnn,” ... for our variables (instead of “x,” “y,” “z,” “x’,” ...). We’ll drop “∃,” and write “~(n)~” instead of “(∃n).” Our minimal-vocabulary version of AC uses only eight symbols:

/   +   •   (   )   =   ~   n

Any statement of arithmetic can be symbolized by combining these symbols.

Our strategy for proving Gödel’s theorem goes as follows. First we give ID numbers to AC formulas. Then we see how AC formulas can encode messages about other AC formulas. Then we construct a special formula, called the Gödel formula G, that encodes this message about itself: “G isn’t provable.” G asserts its own unprovability; this is the key to Gödel’s theorem.

It’s easy to give ID numbers to AC formulas. Let’s assign to each of the eight symbols a digit (an ID number) from 1 to 8:

/	+	•	(	)	=	~	n
1	2	3	4	5	6	7	8

Thus “/” has ID # 1 and “+” has ID # 2. To get the ID number for a formula, we replace each symbol by its one-digit ID number. So we replace “/” by “1,” “+” by “2,” and so on. Here are two examples:

/=/
161

(//+//)
4112115

The ID numbers follow patterns. For example, each numeral has an ID number consisting of all 1’s: 0346

/	//	///	////
1	11	111	1111

So we can say:

Formula # n is a numeral *if and only* if **n consists of all 1’s.**

We can express the right side as the equation “(nine-times-n plus one) equals some power of ten,” or “$(\exists x)9n+1=10^x$,” which can be symbolized in an AC formula.[1] This AC formula is true of any number n if and only if formula # n is a numeral. This is how system AC encodes messages about itself.

An AC theorem is any formula provable in AC. The ID numbers for theorems follow definite but complex patterns. It’s possible to find an equation that’s true of any number n if and only if formula # n is a theorem. If we let “n is …” represent this equation, we can say:

Formula # n is a theorem *if and only if* **n is ….**

The equation on the right side would be very complicated.

To make things more intuitive, let’s pretend that all and only theorems have *odd* ID numbers. Then “n is odd” encodes “Formula # n is a theorem”:

Formula # n is a theorem *if and only if* **n is odd.**

[1] The AC formula for this equation is “~(nn)~(((///////// • n) + /) = (////////// •• nn)).” This formula has ID # 7488574441111111113852156411111111113388855. It’s important that our right-hand bold formulas can be symbolized in AC formulas with definite ID numbers. It’s not important that we write out the formulas or their ID numbers.

For example, “161 is odd” encodes the message that formula # 161 (which is “/=/”) is a theorem:

Formula # 161 is a theorem *if and only if* **161 is odd.**

Then “n is even” would encode the message that formula # n is a non-theorem:

Formula # n is a non-theorem *if and only if* **n is even.**

Imagine that “485…” is some specific very large number. Let “485… is even” represent the AC formula that says that 485… is even:

485… is even.

This formula would encode the following message:

Formula # 485… is a non-theorem. 0347

So the AC formula is true if and only if formula # 485… is a non-theorem. Now suppose this formula itself happens to have ID number 485…. Then the formula would talk about itself, declaring that it itself is a non-theorem. This is what the Gödel formula G does. G, with a certain ID number, encodes the message that the formula with this ID number is a non-theorem. G in effect says this:

G G is not a theorem.

So G encodes the message “G is not a theorem.” But this means that G is true if and only if it’s not a theorem.

So G is true if and only if it’s not provable. Now G, as a formula of arithmetic, is either true or false. Is G true? Then it’s not provable – and our system contains unprovable truths. Or maybe G is false? Then it’s provable – and our system contains provable falsehoods. In either case, system AC is flawed.

We can’t remove the flaw by adding further axioms or inference rules. No matter what we add to the arithmetic calculus, we can use Gödel’s technique to find a formula of the system that’s true-but-unprovable or false-but-provable. Hence arithmetic can’t be reduced to any sound and complete formal system.

This completes our sketch of the reasoning behind Gödel’s proof. To fill in the details would require answering two further questions:

- Consider the equation that's true of any number n if and only if formula # n is a theorem. This equation would have to be much more complicated than "n is odd." How can we produce this equation?
- If we have this equation, how do we then produce a formula with a given number that says that the formula with that number is a non-theorem?

The answers to these questions are too complicated to go into here. While the details can be worked out, we won't here worry about how to do this.[1]

Most people find the last two chapters surprising. We tend to think that *everything* can be proved in math, and that *nothing* can be proved in ethics. But Gödel's theorem shows that not everything can be proved in math. And our golden-rule formalization shows that some important ideas (like the golden rule) can be proved in ethics. Logic can surprise us.

[1] My *Gödel's Theorem Simplified* (Langham, Md.: University Press of America, 1984) has details.

16 History of Logic

Logic was born in ancient Greece and reborn a century ago. Logic keeps growing and expanding, and has contributed to the birth of the computer age. We can better understand and appreciate logic by studying its history.

16.1 Ancient logic

The formal study of valid reasoning began with Aristotle (384–322 BC) in ancient Greece. An unprecedented emphasis on reasoning prepared for Aristotle's logic. Greeks used complex reasoning in geometry, to prove results like the Pythagorean theorem. Sophists taught rich young men to gain power by arguing effectively (and often by verbal trickery). Parmenides and Heraclitus reasoned about being and non-being, anticipating later disputes about the law of non-contradiction, and Zeno reasoned about paradoxes. Socrates and Plato gave models of careful philosophical reasoning; they tried to derive absurdities from proposed views and sought beliefs that could be held consistently after careful examination.

Reasoning is an important human activity, and it didn't begin in ancient Greece. Is this ability biologically based, built into our brains by evolution because it aids survival? Or does it have a divine origin, since we're made in the "image and likeness" of God? Or do both explanations have a place? Logic raises fascinating issues for other disciplines.

Aristotle began the *study* of logic. He was the first to formulate a correct principle of inference, to use letters for terms, and to construct an axiomatic system. He created syllogistic logic (Chapter 2), which studies arguments like these (using "all A is B," "no A is B," "some A is B," or "some A is not B"):

All humans are mortal.
All Greeks are humans.
∴ All Greeks are mortal.

all H is M **Valid**
all G is H
∴ all G is M

This is *valid* because of its formal structure, as given by the formulation on the right; any argument having this same structure will be valid. If we change the structure, we may get an invalid argument, like this one: 0349

All Romans are mortal.
All Greeks are mortal.
∴ All Greeks are Romans.

all R is M **Invalid**
all G is M
∴ all G is R

This is *invalid* because its form is wrong. Aristotle defended valid forms by deriving them from self-evidently valid forms; he criticized invalid forms by showing that they sometimes give true premises and a false conclusion. His logic of syllogisms is about *logic in a narrow sense*, since it deals with what follows from what. He also pursued other topics that connect with appraising arguments, such as definitions and fallacies; these are about *logic in a broader sense*.

Aristotle proposed two principles of thought. His **law of non-contradiction** states that the same property cannot at the same time both belong and not belong to the same object in the same respect. So "S is P" and "S is not P" can't both be true at the same time, unless we take "S" or "P" differently in the two statements. Aristotle saw this law as so certain that it can't be proved by anything more certain; not all knowledge can be demonstrated, since otherwise we'd need an infinite series of arguments that prove every premise by a further argument. Deniers of the law of non-contradiction assume it in their practice; to drive this point home, we might bombard them with contradictions until they plead for us to stop. Aristotle also supported the **law of excluded middle**, that either "S is P" or "S is not P" is true. Some deviant logics today dispute both laws (Chapter 17).

Aristotle also studied the logic of "necessary" and "possible" (see *modal logic*, Chapters 10 and 11). He discussed future contingents (events that may or may not happen). Consider a possible sea battle tomorrow. If "There *will* be a sea battle tomorrow" ("S" below) is *now* either true or false, this seems to make necessary whether the battle occurs:

Either it's true that S or it's false that S.
If it's true that S, then it's necessary that S.
If it's false that S, then it's necessary that not-S.
∴ Either it's necessary that S or it's necessary that not-S.

Aristotle rejected the conclusion, saying that there was no necessity either way. He seemed to deny the first premise and thus the universal truth of the law of excluded middle (which he elsewhere defends); if we interpret him this way, then he anticipated many-valued logic in using a third truth value besides *true* and *false* (§17.1). Another solution is possible. Many think premises 2 and 3 have a box-inside/box-outside ambiguity (§10.1): taking them as "$(A \supset \Box B)$" makes them doubtful while taking them as "$\Box(A \supset B)$" makes the argument invalid.

After Aristotle, Stoics and others developed a logic that focused on "if-then," "and," and "or," like our propositional logic (Chapters 6 and 7). Stoic logicians defended, for example, an important inference form that came to be called *modus tollens* (denying mode): 0350

If your view is correct, then such and such is true.
Such and such is false.
∴ Your view isn't correct.

If C then S **Valid**
Not-S
∴ Not-C

Stoics also studied modal logic. Unlike logicians today, they took "necessary" and "possible" in a temporal sense, like "true at all times" and "true at some times." They disputed whether there was a good modal argument for *fatalism*, the view that all events happen of inherent necessity (see §10.3b #10). They also disputed how to understand "If A then B" (§17.4). Philo of Megara saw it as true if and only if it's not *now* the case that A is true and B is false; this fits the modern truth table for "if-then." Diodorus Chronos saw it as true if and only if A is *never* at any time true while B is false.

Aristotelian and Stoic logic were first seen as rivals, differing in three ways:

- Aristotle focused on "all," "no," and "some." Stoics focused on "if-then," "and," and "or."
- Aristotle used letter variables and expressed arguments as long conditionals, like "If all A is B, and all C is A, then all C is B." Stoics used number variables and expressed arguments as sets of statements, like "If 1 then 2. But not-2. Therefore, not-1."
- Aristotle saw logic not as part of philosophy but rather as a tool for all think-

> ing. Stoics saw logic as one of philosophy's three branches (the other two being physics and ethics). But both agreed that students should study logic early, before going deeply into other areas.

Later thinkers combined these approaches into **traditional logic**. For the next two thousand years, Aristotle's logic with Stoic additions ruled in the West.

At the same time, another tradition of logic rose up in India, China, and Tibet. We call it **Buddhist logic** even though Hindus and others pursued it too. It studied many topics important in the West, including inference, fallacies, and language. This is a common pattern in Buddhist logic:

Here there is fire, because there is smoke.
Wherever there is smoke there is fire, as in a kitchen.
Here there is smoke.
∴ Here there is fire.

The last three lines are deductively valid:

All cases of smoke are cases of fire.
This is a case of smoke.
∴ This is a case of fire.

This omits "as in a kitchen," which suggests inductive reasoning (Chapter 5); in our experience of smoke and fire, smoke always seems to involve fire. 0351

The Eastern logic tradition is poorly understood in the West; this tradition covers many centuries, and many texts are difficult or untranslated. Some commentators emphasize similarities between East and West; they see human thinking as essentially the same everywhere. Others emphasize differences and caution against imposing a Western framework on Eastern thought. And some deviant logicians see the Eastern tradition as congenial to their views.

Many see the East as more mystical than logical; Zen Buddhism delights in using paradoxes (like the sound of one hand clapping) to move us beyond logical thinking toward a mystical enlightenment. But East and West both have logical and mystical elements. Sometimes these come together in the same individual; Ludwig Wittgenstein in the early 20th century invented truth tables but also had a strongly mystical side.

16.2 Medieval logic

Medieval logicians carried on the basic framework of Aristotle and the Stoics, as logic became important in higher education.

The Christian thinker Boethius (480–524) helped the transition to the Middle Ages. He wrote on logic, including commentaries; he explained the modal box-inside/box-outside ambiguity as he defended the compatibility of divine foreknowledge and human freedom (§10.3b #4 and #14). He translated Aristotle's logic into Latin. Many of his translations were lost; but his *Categories* and *On Interpretation* became the main source for the *logica vetus* (old logic).

The Arab world dominated in logic from 800–1200. Some Arab logicians were Christian, but most were Muslim; both groups saw logic as important for theology and medicine. They translated Aristotle into Arabic and wrote commentaries, textbooks, and original works. They pursued topics like syllogisms, modal logic, conditionals, universals, predication, and existence. Baghdad and Moorish Spain were centers of logic studies.

In Christian Europe, logic was reborn in the 11th and 12th centuries, with Anselm, Peter Abelard, and Latin translations of Aristotle's *Prior Analytics*, *Posterior Analytics*, *Topics*, and *Sophistical Refutations*; the *logica nova* (new logic) was based on these. There was interest in universals and in how terms signify. Peter of Spain and William of Sherwood wrote logic textbooks.

The clever Barbara-Celarent verse was a tool for teaching syllogisms:

Barbara, Celarent, Darii, Ferioque, prioris;
Cesare, Camestres, Festino, Baroco, secundae;
tertia, Darapti, Disamis, Datisi, Felapton,
Bocardo, Ferison, habet; quarta insuper addit
Bramantip, Camenes, Dimaris, Fesapo, Fresison.

Capitalized names are valid syllogisms. Vowels are sentence forms: 0352

A = "all … is …"
I = "some … is …"

Aff-**I**rm universal/particular

E = "no … is …"
O = "some … is not …"

n**E**-g**O** universal/particular

So "Barbara," with AAA vowels, has three "all" statements:

all M is P
all S is M
∴ all S is P

Figure 1 (MP / SM in premises)

Aristotelian syllogisms have two premises. *Middle term* "M" is common to both premises; *predicate* "P" occurs in the first premise, while *subject* "S" occurs in the second. There are four figures (arrangements of premise letters):

1	2	3	4
prioris	*secundae*	*tertia*	*quarta*
MP SM	PM SM	MP MS	PM MS

Aristotle's four axioms are valid first-figure forms:

Barbara

all M is P
all S is M
∴ all S is P

Celarent

no M is P
all S is M
∴ no S is P

Ferio

all M is P
some S is M
∴ some S is P

Darii

no M is P
some S is M
∴ some S is not P

The other 15 forms can be derived as theorems. The consonants give clues on how to do this; for example, "m" says to switch the order of the premises.

Thomas Aquinas (1224–74), the most influential medieval philosopher, had little impact on logic's development; but he made much use of logic. Since he emphasized reasoning and wrote so much, he likely produced more philosophical arguments than anyone else who has ever lived.

Fourteenth-century logicians include William of Ockham and Jean Buridan. *Ockham's razor* says "Accept the simplest theory that adequately explains the data." Ockham developed modal logic and tried to avoid metaphysics when analyzing language. *Buridan's ass* was a fictional donkey whose action was paralyzed when he was placed exactly midway between two food bowls. Buridan also formulated the standard rules for valid syllogisms; one version says that a syllogism is *valid* just if it satisfies all of these conditions:

- Every term distributed in the conclusion must be distributed in the premises. (A term is *distributed* in a statement just if the statement makes some claim about *every* entity that the term refers to.)
- The middle term must be distributed in at least one premise. (The *middle term* is the one common to both premises; if we violate this rule, we commit the 0353 fallacy of the *undistributed middle*.)
- If the conclusion is negative, exactly one premise must be negative. (A statement is *negative* if it contains "no" or "not"; otherwise it's positive.)
- If the conclusion is positive, both premises must be positive.

In the Middle Ages, logic was important in philosophy and in higher education. Even today, logic, like biology, uses many Latin terms (*modus ponens*, *a priori*/*a posteriori*, *de re*/*de dicto*, and so on).

16.3 Enlightenment logic

Aristotelian logic dominated until the end of the 19th century. Several logicians contributed to syllogistic logic; for example, Leonhard Euler diagrammed "all A is B" by putting an A-circle inside a larger B-circle, Lewis Carroll entertained us with silly syllogisms and points about logic in *Alice in Wonderland*, and John Venn gave us diagrams for testing syllogisms (§2.6). But most logicians would have agreed with Immanuel Kant, who said that Aristotle invented and perfected logic; nothing else of fundamental importance could be added, although we might improve teaching techniques. Kant would have been shocked to learn about the revolution in logic that came about a hundred years later.

The German thinkers Georg Hegel and Karl Marx provided a side current. Hegel proposed that logic should see contradictions as explaining how thought evolves historically; one view provokes its opposite, and then the two come together in a higher synthesis. Marx saw contradictions in the world as real; he applied this to political struggles and revolution. While some saw this *dialectical logic* as an alternative to traditional logic, critics objected that this confuses conflicting properties in the world (like hot/cold or capitalist/proletariat) with logical self-contradictions (like the same object being both white and, in the same sense and time and respect, also non-white).

The philosopher Gottfried Leibniz, the co-inventor of calculus, anticipated future developments. He proposed the idea of a symbolic language that would reduce reasoning to calculation. If controversies arose, the parties could take up their pencils and say, "Let us calculate." Leibniz created a logical notation much like that of Boole (and much earlier); but his work was published after Boole.

Many thinkers tried to invent an algebraic notation for logic. Augustus De Morgan proposed symbolizing "all A is B" as "A))B" and "some A is B" as "A()B"; a letter on the concave side of the parenthesis is distributed. He became known for his *De Morgan laws* for propositional logic:

Not both A and B	=	Either not-A or not-B
Not either A or B	=	Both not-A and not-B

He complained that current logic couldn't handle relational arguments like "All 0354 dogs are animals; therefore all heads of dogs are heads of animals" (§9.5b #25).

The **Boolean algebra** of George Boole (1815–64) was a breakthrough, since it used math to check the correctness of inferences. Boole used letters for sets; so "M" might be the set of mortals and "H" the set of humans. Putting two letters together represents the *intersection* of the sets; so "HM" is the set of those who are *both human and mortal*. Then "All humans are mortal" is "H = HM," which says that the set of humans = the set of those who are both human and mortal. A syllogism is a series of equations:

All humans are mortal.
All Greeks are humans.
∴ All Greeks are mortal.

H = HM   **Valid**
G = GH
∴ G = GM

We can derive the conclusion by substituting equals for equals. In premise 2, G = GH, replace "H" with "HM" (premise 1 says H = HM) to get G = GHM. Then replace "GH" with "G" (premise 2 says G = GH) to get G = GM.

Boolean formulas, like those below (which use a later symbolism), can be interpreted to be about sets or about statements:

"-A" can mean "the set of non-As" or "not-A"

"A∩B" can mean "the intersection of sets A and B" or "A and B"

"A∪B" can mean "the union of sets A and B" or "A or B"

So if "A" is the set of animals, then "-A" is the set of non-animals; but if "A" is "Aristotle is a logician," then "-A" is "Aristotle isn't a logician." The same laws cover both; for example, "A∩B = B∩A" works for either sets or statements. *Boolean operators* (like "and," "or," and "not") use the statement interpretation.

Boole, the father of *mathematical logic*, thought that logic belonged with mathematicians instead of philosophers. But both groups came to have an interest in logic, each getting the slice of the action that fits it better. While Boole was important, a greater revolution in logic was to come.

16.4 Frege and Russell

Gottlob Frege (1848–1925) created modern logic with his 1879 *Begriffsschrift* ("Concept Writing"). Its 88 pages introduced a symbolism that, for the first time, let us combine in every way Aristotle's "all," "no," and "some" with the Stoic "if-then," "and," and "or." So we can symbolize "If everything that's A or B is then C and D, then everything that's non-D is non-A." Thus the gap between Aristotle and the Stoics was overcome in a higher synthesis. Frege also showed how to analyze arguments with relations (like "x loves y") and multiple quantifiers; so we can show that "There is someone that everyone loves" entails "Everyone loves someone" – but not conversely. Frege presented logic as a 0355 *formal system*, with purely notational rules for determining the grammaticality of formulas and the correctness of proofs.

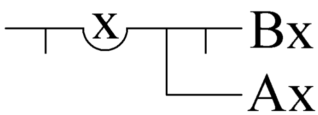
Frege's work was ignored until Bertrand Russell (1872–1970) praised it in the early 20th century. Frege's difficult symbolism alienated people. He used lines for "not," "if-then," and "all":

Not-A

If A then B

For all x

These can combine to symbolize “Not all A is non-B” (our “∼(x)(Ax ⊃ ∼Bx)”):



This was also his way to write “Some A is B” (our “(∃x)(Ax • Bx)”); he had no simpler notation for “some” or “and.”

Frege developed logic to help show that arithmetic is reducible to logic; he wanted to define all basic concepts of arithmetic (like numbers and addition) in purely logical terms and prove all basic truths of arithmetic using just logical axioms and inference rules. Frege used a seemingly harmless axiom that every condition on x picks out a set containing just those elements that satisfy that condition; so the condition “x is a cat” picks out the set of cats. But consider that some sets are members of themselves (the set of abstract objects is an abstract object) while other sets aren’t (the set of cats isn’t a cat). By Frege’s axiom, “x is not a member of itself” picks out the set containing just those things that are not members of themselves. Call this “set R.” So any x is a member of R, if and only if x is not a member of x (here “∈” means “is a member of” and “∉” means “is not a member of”):

For all x, x ∈ R if and only if x ∉ x.

Russell asked in a 1902 letter to Frege: What about set R itself? By the above principle, R is a member of R, if and only if R is not a member of R:

R ∈ R if and only if R ∉ R.

So is R a member of itself? If it is, then it isn’t – and if it isn’t, then it is; either way we get a contradiction. Since this contradiction, called **Russell’s paradox**, was provable in Frege’s system, that system was flawed. Frege was crushed, since his life work collapsed. His attempts to fix the problem weren’t successful.

Russell greatly admired Frege and his groundbreaking work in logic; the two minds worked along similar lines. But the paradox showed that Frege’s work needed fixing. So Russell, with his former teacher Alfred North Whitehead, 0356 worked to develop logic and set theory in a way that avoided the contradiction. They also developed a more intuitive symbolism (much like what we use in this book), based on the work of Giuseppe Peano. The result was their massive *Principia Mathematica*, which was published in 1910–1913. *Principia* had a huge influence and became the standard formulation of the new logic.

16.5 After *Principia*

Classical symbolic logic includes propositional and quantificational logic (Chapters 6 to 9). A logic is "classical" if it accords with Frege and Russell about which arguments are valid, regardless of differences in symbolization and proof techniques. Classical symbolic logic gradually became the new orthodoxy, replacing the older Aristotelian logic.

Much work was done to solidify classical symbolic logic. Different proof techniques were developed; while Frege and Russell used an axiomatic approach, later logicians invented inferential and truth-tree methods that were easier to use. Different ways of symbolizing arguments were developed, including the *Polish notation* of a school of logic that was strong in Poland between the world wars. Ludwig Wittgenstein and Emil Post independently invented truth tables, which clarified our understanding of logical connectives (like "if-then," "and," and "or") and led to a criterion of validity based on semantics – on the meaning of the connectives and how they contribute to truth or falsity; Alfred Tarski and others expanded the semantic approach to quantificational logic.

Much work was done in **metalogic**, the study of logical systems (Chapter 15). Kurt Gödel showed that Russell's axiomatization of classical logic was, given certain semantic assumptions, correct: just the right things were provable. But he also showed, against Frege and Russell, that arithmetic cannot be reduced to any formal system: no consistent set of axioms and inference rules would suffice to prove all arithmetic truths; this result, called **Gödel's theorem**, is perhaps the most striking and surprising result of 20th-century logic. Alonzo Church showed that the problem of determining validity in quantificational logic cannot be reduced to an mechanical algorithm (a result called *Church's theorem*). There was also much activity in *set theory*, which after Russell's paradox became increasingly complex and controversial.

There was also much work in **philosophy of logic** (Chapter 18), which deals with philosophical questions about logic, such as these: Are logical truths dependent on human conventions (so different conventions might produce different logical truths) or on the objective nature of reality (perhaps giving us the framework of any possible language that would be adequate to describe reality)? Can logic help us clarify metaphysical issues, such as what kinds of entity ultimately exist? Should we assume abstract entities (like properties and propositions) when we do logic? How can we resolve logical paradoxes (such as Russell's 0357 paradox and the liar paradox)? Are logical truths empirical or *a priori*? Does logic distort ordinary beliefs and ordinary language, or does it correct them? What is the definition and scope of logic?

Logic was important in the development of computers. The key insight here was that logical functions like "and" and "or" can be simulated electrically by *logic gates*; this idea goes back to the American logician Charles Sanders Peirce in the 1880s and was rediscovered by Claude Shannon in 1938. A computer contains logic gates, plus memory and input-output devices. Logicians like John von Neumann, Alan Turing, and Arthur Burks helped design the first large-scale electronic computers. Since logic is important for computers, in both hardware and software, it's studied today in computer science departments. So now three main departments study logic – philosophy, math, and computer science.

Logic today is also an important part of *cognitive science*, an interdisciplinary approach to thought that includes linguistics, psychology, biology (brain and sensory systems), computers (especially artificial intelligence), and other branches of philosophy (especially epistemology and philosophy of mind).

As classical symbolic logic became the orthodoxy, it started to be questioned. Two types of non-classical logic came to be. *Supplementary logics* accepted that classical logic was fine as far as it went but needed to be supplemented to deal, for example, with "necessary" and "possible." **Deviant logics** thought that classical logic was wrong on some points and needed to be changed.

The most important *supplementary logic* is modal logic, which deals with "necessary" and "possible" (Chapters 10 and 11). Ancient and medieval logicians pursued modal logic; but 20th-century logicians mostly ignored it until C. I. Lewis's work in the 1930s. Modal logic then became controversial. Willard Quine argued that it was based on a confusion; he thought logical necessity was unclear and quantified modal logic led to an objectionable metaphysics of necessary properties. There was lively debate on modal logic for many years. In 1959, Saul Kripke presented a possible-worlds way to explain modal logic; this made more sense of it and gave it new respect among logicians. Possible worlds have proved useful in other areas and are now a common tool in logic; and several philosophers (including Alvin Plantinga) have defended a metaphysics of necessary properties. Today, modal logic is a well-established extension of classical logic.

Other extensions apply to ethics ("A ought to be done" or "A is good"), theory of knowledge ("X believes that A" or "X knows that A"), the part–whole relationship ("X is a part of Y"), temporal relationships ("It will be true at some future time that A" and "It was true at some past time that A"), and other areas (Chapters 12 to 14). Most logicians would agree that classical logic needs to be supplemented in order to cover certain kinds of argument.

Deviant logics say that classical symbolic logic is wrong on some points and needs to be changed (Chapter 17). Some propose using more than two truth values. Maybe we need a third truth value for "half-true." Or maybe we need a fuzzy-logic range of truth values, from completely true (1.00) to completely false (0.00). Or perhaps "A" and "not-A" can both be false (intuitionist logic) or both 0358 be true (paraconsistent logic). Or perhaps the classical approach to "if-then" is flawed; some views even reject *modus ponens* ("If A then B, A ∴ B") and *modus tollens* ("If A then B, not-B ∴ not-A"). These and other deviant logics have been proposed. Today there is much questioning of basic logical principles.

This brief history of logic has focused on deductive logic and related areas. There has also been much interest in informal logic (Chapters 3 and 4), inductive logic (Chapter 5), and history of logic (this chapter).

So logic has a complex history – from Aristotle and the Stoics in ancient Greece, through the Middle Ages and the Enlightenment, to the turmoil of the 19th century and logic's transformation with Frege and Russell, and into recent classical and non-classical logics and the birth of the computer age.[1]

[1] For more on the history of logic, I suggest P. H. Nidditch's *The Development of Mathematical Logic* (London: Routledge & Kegan Paul, 1962) and, for primary sources, Irving Copi and James Gould's *Readings on Logic* (New York: Macmillan, 1964). Also useful are William and Martha Kneale's *The Development of Logic* (Oxford: Clarendon, 1962) and Joseph Bocheński's *A History of Formal Logic*, trans. Ivo Thomas (Notre Dame, Ind.: University of Notre Dame, 1961).

17 Deviant Logics

Deviant logics reject standard assumptions. Most logicians have assumed that statements are either true or false, but not both, and that *true* and *false* are the only truth values. Deviant logics question such ideas. Maybe we need more than two truth values (many-valued logic). Or maybe "A" and "not-A" can both be true (paraconsistent logic) or both be false (intuitionist logic). Or maybe standard IF-THEN inferences are mistaken (relevance logic).

Deviant logics are controversial. Some are happy that logic is becoming, in some circles, as controversial as other areas of philosophy. Others defend standard logic and see deviant logic as promoting intellectual chaos; they fear what would happen if thinkers couldn't assume that *modus ponens* and *modus tollens* are valid and that contradictions are to be avoided.

17.1 Many-valued logic

Most logicians assume that there are only two truth values: *true* and *false*. Our propositional logic in Chapter 6 accepts this **bivalence**, symbolizing true as "1" and false as "0." This is consistent with there being truth-value gaps for sentences that are meaningless (like "Glurklies glurkle") or vague (like "Her shirt is white," when it's between white and gray). Logic needn't worry about such sentences, since arguments using them are already defective; so we can just stipulate that capital letters stand for statements that *are* true or false.

Many-valued logics accept more than two truth values. **Three-valued logic** might use "1" for true, "0" for false, and "½" for *half-true*. This last category might apply to statements that are unknowable, or too vague to be true-or-false, or plausible but unproved, or meaningless, or about future events not yet decided. A three-valued truth table for NOT looks like this:

P	~P
0	1
½	½
1	0

If P is false, then ~P is true.
If P is half-true, then ~P is half-true.

If P is true, then ~P is false.

This table shows how the other connectives work: 0360

P	Q	(P • Q)	(P ∨ Q)	(P ⊃ Q)	(P ≡ Q)
0	0	0	0	1	1
0	½	0	½	1	½
0	1	0	1	1	0
½	0	0	½	½	½
½	½	½	½	1	1
½	1	½	1	1	½
1	0	0	1	0	0
1	½	½	1	½	½
1	1	1	1	1	1

AND takes the value of the lower conjunct, and OR takes the value of the higher disjunct. IF-THEN is true if the consequent is at least as true as the antecedent and is half-true if its truth is a little less. IF-AND-ONLY-IF is true if both parts have the same truth value and is half-true if they differ a little.

Given these truth tables, some standard logical laws fail. "(P ∨ ~P)" (the law of excluded middle) and "~(P • ~P)" (the law of non-contradiction) sometimes are only half true. "(P ⊃ Q)" isn't equivalent to "~(P • ~Q)," since they differ in truth value if P and Q are both ½. We can avoid these results by making "(½ ∨ ½)" true and "(½ • ½)" false; but then "P" strangely isn't logically equivalent to "(P ∨ P)" or "(P • P)."

Fuzzy logic proposes an infinity of truth values; these can be represented by real numbers between 0.00 (fully false) and 1.00 (fully true). We might define a "valid argument" as one in which, if the premises have at least a certain truth value (perhaps .9), then so does the conclusion; *modus ponens* then fails (since if "A" and "(A ⊃ B)" are both .9, then "B" might be less than .9) as do some other logical principles. Some propose an even fuzzier logic with vaguer truth values like "very true" or "slightly true."

Fuzzy logic is used in devices like clothes dryers to permit precise control. A crisp-logic dryer might have a rule that if the shirts are dry then the heat is turned off; a fuzzy-logic dryer might say that if the shirts are dry to degree n then the heat is turned down to degree n. We could get the same result using standard logic and a relation "Dxn" that means "shirt x is dry to degree n" – thus using degrees-of-dryness instead of degrees-of-truth.

Opponents say many-valued logic is weird and arbitrary and has little application to real-life arguments. Even if this is so, the many-valued approach has other applications. It can be used, for example, in computer memory systems with more than two states. And it can be used to show the independence of axioms for propositional logic (§15.5); an axiom can be shown to be independent of the other axioms of a certain system if, for example, the other axioms (and theorems derived from these) always have a value of "7" on a given truth-table scheme, while this axiom sometimes has a value of "6." 0361

17.2 Paraconsistent logic

Aristotle's **law of non-contradiction** states that the same property cannot at the same time both belong and not belong to the same object in the same respect. So "S is P" and "S is not P" cannot both be true at the same time, unless we take "S" or "P" differently in the two statements. Aristotle saw this law as certain but unprovable. Deniers of the law assume it in their practice; wouldn't they complain if we bombarded them with contradictions?

Aristotle mentions Heraclitus as denying the law of non-contradiction. The 19th-century thinkers Georg Hegel and Karl Marx also seemed to deny it and are often seen as proposing an alternative *dialectical logic* in which contradictions are real. Critics object that such a logic would confuse conflicting properties in the world (like hot/cold or capitalist/proletariat) with logical self-contradictions (like the same object being both white and, in the same sense and time and respect, also non-white).

In standard propositional logic, the law of non-contradiction is "∼(P • ∼P)" and is a *truth-table tautology* – a formula true in all possible cases:

P	∼(P • ∼P)
0	1
1	1

"This is false: I went to Paris and I didn't go to Paris."

"P and not-P" is always false in standard logic, which presupposes that "P" stands for the same statement throughout. English is looser and lets us shift the meaning of a phrase in the middle of a sentence. "I went to Paris and I didn't go to Paris" may express a truth if it means "I went to Paris (in that I landed once at the Paris airport) – but I didn't really go there (in that I saw almost nothing of the city)." Because of the shift in meaning, this would better translate as "(P • ∼Q)," which wouldn't violate the law of non-contradiction.

Some recent logicians, like Graham Priest, claim that *sometimes* a statement and its contradictory are both true. Such **dialethist** logicians don't say that *all* statements and their denials are true – but just that *some* are. Here are examples where "A and not-A" might be claimed to be true:

- "We do and don't step into the same river." (Heraclitus)

- “God is spirit and isn’t spirit.” (the mystic Pseudo-Dionysius)
- “The moving ball is here and not here.” (Hegel and Marx)
- “The round square is both round and not-round.” (Meinong)
- “The one hand claps and doesn’t clap.” (Eastern paradox)
- “Sara is a child and not a child.” (paradoxical speech)
- “What I am telling you now is false.” (liar paradox)
- “The electron did and didn’t go in the hole.” (quantum physics)

Most logicians contend that these aren’t genuine cases of “A and not-A,” at least 0362 if they’re taken in a sensible way, since we must take the two instances of “A” to represent different ideas. For example, “Sara is a child and not a child” can be sensible only if it really means something like “Sara is a *child-in-age* but not a *child-in-sophistication*.” Paradoxical speech, although sometimes nicely provocative, doesn’t make sense if taken literally. Dialethists try to show that some of their allegedly true self-contradictions resist such analyses.

In standard propositional logic we can from a single self-contradiction deduce the truth of every statement and its denial. But then, if we believed a self-contradiction and also all its logical consequences, we’d contract the dreaded disease of *contradictitis* – whereby we’d believe every statement and also its contradictory – bringing chaos to human speech and thought. Here’s an intuitive derivation showing how, given the contradictory premises “A is true” and “A is not true,” we can deduce any arbitrary statement “B” (this “A, ∼A ∴ B” inference is called the **explosion principle**):

1 A is true. {premise}
2 A is not true. {premise}
3 ∴ At least one of these two is true: A or B. {from 1: if A is true then at least one of the two, A or B, is true}
4 ∴ B is true. {from 2 and 3: if at least one of the two, A or B, is true and it’s not A, then it’s B}

Dialethists respond by rejecting standard logic. They defend a **paraconsistent logic** that rejects the explosion principle; this lets them contain an occasional self-contradiction without leading to an “anything goes” logical nihilism. In the above argument, they reject line 4 and thus the “(A ∨ B), ∼A ∴ B” inference (*disjunctive syllogism*). Suppose, they say, B is false and A is both-true-and-false (!); then, they say, “(A ∨ B)” is true (since “A” is true), “∼A” is true (since “A” is also false), but “B” is false – and so disjunctive syllogism is invalid.

Paraconsistent logicians have developed their own truth tables. One option lets “A” and “not-A” have any combination of *true* or *false*, independently of each other; so we have four possibilities:

P	~P
0	0
0	1
1	0
1	1

P and not-P are both false.
P is false and not-P is true.
P is true and not-P is false.
P and not-P are both true.

This approach rejects the usual understanding of "not," whereby "not-A" has the opposite truth value as "A." In paraconsistent logic, disjunctive syllogism is invalid, since it can have true premises and a false conclusion:

A	~A	B	(A ∨ B),	~A	∴	B
1	1	0	**1**	**1**		**0**

0363 Similarly, the explosion principle, which permits us to deduce any arbitrary statement from a self-contradiction, is invalid:

A	~A	B	A,	~A	∴	B
1	1	0	**1**	**1**		**0**

Paraconsistent logic lets logic go on normally for the most part – so most of the arguments in this book that are valid/invalid on standard logic would come out the same as before; but it also permits an occasional self-contradiction to be true. Thus it denies that a strict adherence to the law of non-contradiction is necessary for coherent thought.

Critics object that it makes no sense to permit "A" and "not-A" to both be true, at least if we take "not" in anything close to its normal sense. If we reject the usual truth table for "not," which makes "not-A" always have the opposite truth value of "A," then what is left of the meaning of "not"?

Critics also object that permitting "A" and "not-A" to both be true lets irrational people off too easily. Imagine politicians or students who regularly contradict themselves, asserting "A" and then a few breaths later asserting "not-A," and yet defend themselves using the "new logic," which lets both be true at once. Surely this is lame and sophistical.

Some who accept the law of non-contradiction see value in paraconsistent logic, since people or computers may have to derive conclusions from inconsistent data. Suppose that our best data about a crime is flawed and inconsistent; we still might want to derive the best conclusions we can from this data. The “anything and its opposite follows from inconsistent data” approach of classical logic is unhelpful. Paraconsistent logic claims to do better.

Critics question whether paraconsistent logic *can* do better. If our data is inconsistent, then it has errors and can’t provide reliable conclusions. So we need to clear up the inconsistency first, perhaps by rejecting the least solidly based statements. We need to see what follows (using standard logic) from the most probable consistent subset of the original data.

Critics also claim that rejecting disjunctive reasoning lessens the real-world usefulness of paraconsistent logic. Suppose we know that either A or B committed the murder, and then we find out that A didn’t do it. We need to conclude that B then did it; but paraconsistent logic says that this is invalid!

Logicians defend the law of non-contradiction in different ways. Some see it as a useful language convention. We could imagine a tribe where vague statements (like “This shirt is white”) in borderline cases are said to be *both true and false* (instead of *neither true nor false*). Some might speak this way; and we could easily translate between this and normal speech. If so, then perhaps a strict adherence to the law of non-contradiction is at least partly conventional. But, even so, it’s a convention that’s less confusing than the paraconsistent alternative.

Others see the law of non-contradiction as a deep metaphysical truth about reality. They see paraconsistent logicians as offering, not an alternative way of 0364 speaking, but rather an incoherent metaphysics. Regardless of our verdict here, dialethism and paraconsistent logic do offer interesting challenges that make us think more deeply about logic.

17.3 Intuitionist logic

Aristotle held the **law of excluded middle**, that either “S is P” or “S is not P” is true. Standard propositional logic expresses this as “$(A \vee \sim A)$” (“A or not-A”), which has an all-1 truth table and is true in all possible cases. **Intuitionist logicians**, like the mathematicians Luitzen Brouwer and Arend Heyting, reject this law when applied to some areas of math. They similarly reject the law of *double negation* “$(\sim\sim A \supset A)$” (“If not-not-A, then A”). They think “A” and “$\sim A$” are sometimes both false in cases involving infinite sets. To emphasize these differences, intuitionists use “$\neg$” for negation instead of “$\sim$.”

Intuitionist mathematicians see the natural numbers (0, 1, 2, . . .) as grounded in our experience of counting. Mathematical formulas are human constructs; they shouldn’t be considered *true* unless the mind can prove their truth. Goldbach’s conjecture says “Every even number is the sum of two primes.” This seems to hold for every even number we pick: 2 (1 + 1), 4 (3 + 1), 6 (5 + 1), 8 (7 + 1), and so on. But no one has proved or disproved that it holds for *all* even numbers. Most think Goldbach’s conjecture must be true or false objectively, even if a proof either way may be impossible. Intuitionists disagree. They say *truth* in mathematics is *provability*; if neither Goldbach's conjecture nor its negation is provable, then neither is true. So intuitionists think that, in some cases involving infinite sets (like the set of even numbers), neither "A" nor "∼A" is true, and so both are false. The law of excluded middle *does* apply if we use *finite* sets; so "Every even number under 1,000,000,000 is the sum of two primes" is true or false, and we could write a computer program that could in principle eventually tell us which it is.

Some non-realists reject the law of excluded middle in other areas. Suppose you think the only basic objective truths are ones about your individual experience, like "I feel warmth" or "I sense redness." You might accept objective truths about material objects (like "I'm holding a red pen"), but only if these can be verified by your experience. But often your experience can verify neither "A" nor "not-A"; then neither would be true, and both would be false. So you might reject the law of excluded middle on the basis of a non-realist metaphysics.

Realists think this is bad metaphysics. Goldbach's conjecture about mathematics is objectively true or false; and our experience supports (but doesn't conclusive prove) that it's true. It's wrong to identity "true" with "verified," since we may imagine unverifiable truths; there may be a whole world of truths and falsehoods that aren't accessible to our finite minds. 0365

17.4 Relevance logic

Classical propositional logic analyzes "If P then Q" as simply denying that we have P-true-and-Q-false:

(P ⊃ Q) = ∼(P • ∼Q)

If P is true, then Q is true = We don't have P true and Q false

An IF-THEN understood this way is a **material implication** and is automatically true if the antecedent is false or the consequent is true. This leads to the so-called *paradoxes of material implication*:

- From "not-A" we can infer "If A then B." So from "Pigs don't fly" we can infer "If pigs fly, then I'm rich."
- From "B" we can infer "If A then B." So from "Pigs don't fly" we can infer "If I'm rich, then pigs don't fly."

While many logicians see such results as odd but harmless, relevance logicians see them as wrong and want to reconstruct logic to avoid them.

Relevance logicians oppose evaluating the truth of "If A then B" just by the truth values of the parts; they say an IF-THEN can be true only if the parts are *relevant* to each other. While they're vague on what this means, they insist that logic shouldn't prove theorems like "If A-and-not-A, then B," where antecedent and consequent share no letters. Since paraconsistent logic (§17.2) rejects the related *explosion principle* that a self-contradiction entails every statement, there's a natural affinity between the approaches; many relevance logics are also paraconsistent. Relevance logics often symbolize *relevant implication* as "→," to contrast with the "⊃" of *material implication*.

Defenders of material implication appeal to *conversational implication* to diffuse objections based on the paradoxes of material implication. Paul Grice claims that what is true may not be sensible to assert in ordinary speech. When we speak, we shouldn't make a weaker claim rather than a stronger one unless we have a special reason. Suppose you tell your five children, "At least three of you will get Christmas presents" – while you know that all five will. The weaker statement suggests or insinuates that not all five will get presents. This is due to speech conventions, not logical entailments. "At least three will get presents" doesn't logically entail "Not all five will get presents"; but saying the first insinuates the second. Similarly, there's little point in saying "If P then Q" on the basis of knowing not-P or knowing Q – since it's better to say straight off that not-P or that Q. There's generally a point to saying "If P then Q" only if there's a special connection between the two, some way of going from one to the other. But, again, this has to do with speech conventions, not with truth conditions for "If P then Q." 0366

Some defenders of material implication claim that the so-called paradoxes of material implication are perfectly correct and can be defended by intuitive arguments. We can derive "If not-A then B" from "A":

1 A is true. (Premise)
2 ∴ Either A is true or B is true. {from 1}
3 ∴ If A isn't true, then B is true. {from 2}

Relevance logic must reject this plausible derivation; it must deny that 2 follows from 1, that 3 follows from 2, or that deducibility is *transitive* (if 3 follows from 2, and 2 from 1, then 3 follows from 1). Doing any of these violates our logical intuitions at least as much as do the material-implication paradoxes. So relevance logics, although they try to avoid unintuitive results about conditionals, cannot achieve this goal; they all result in oddities at least as bad as the ones they're trying to avoid. Another problem is that a wide range of conflicting relevance logics have been proposed; these disagree much on which arguments involving conditionals are valid.

Relevance logicians have found other conditional arguments that, while valid on the traditional view, seem to them to be invalid. Some even question the validity of ***modus ponens*** ("If A then B, A ∴ B"). One allegedly questionable *modus ponens* inference involves measles:

If you have red spots, then you have measles.
You have red spots.
∴ You have measles.

(R ⊃ M)
R
∴ M

This is claimed to be invalid because you might have red spots for some other reason. Another objection, from Vann McGee, is more complex. In 1980, three main candidates ran for US president: two Republicans (Ronald Reagan, who won with over 50 percent of the vote, and John Anderson, who got about 7 percent of the vote and was thought to have no chance to win) and a Democrat (Jimmy Carter, who got just over 40 percent). Consider this argument, given just before the election:

If a Republican will win, then if Reagan does not win then Anderson will win.
A Republican will win.
∴ If Reagan does not win, then Anderson will win.

(W ⊃ (~R ⊃ A))
W
∴ (~R ⊃ A)

Here it seems right to believe the premises but not the conclusion (since clearly if Reagan doesn't win, then Carter will win, not Anderson). Again, this instance of *modus ponens* is claimed to be invalid.

Defenders of *modus ponens* think such examples confuse a genuine IF-THEN with other things. Compare these three ways of taking "If you have red spots, then you have measles": 0367

1. *Genuine IF-THEN*: "If you have red spots, then you have measles."
2. *Conditional Probability*: "The probability is high that you have measles, given that you have red spots."
3. *Qualified IF-THEN*: "If you have red spots and other causes can be excluded, then you have measles."

The premise about measles, if a genuine IF-THEN, has to mean 1, and not 2 or 3; but then its truth excludes your having red spots but no measles. The truth of this IF-THEN doesn't entail that we're *certain* that there are no other causes; but *if in fact there are other causes* (so you have red spots but no measles), then the IF-THEN is false. A similar analysis takes care of the Reagan argument.

Even if we reject relevance logic, still we have to admit that some conditionals, or their near relatives, cannot plausibly be interpreted as material implications. We already mentioned *conversational implication* (where saying A suggests or insinuates a further statement B) and *conditional probability* (where fact A would make fact B probable to a given degree). There are also *logical entailments* ("B logically follows from A" – which Chapter 10 symbolizes as "$\Box(A \supset B)$") and *counterfactuals* ("If A had happened then B would have happened" – often symbolized as "$(A \Box\!\!\rightarrow B)$"). So conditionals and their near relatives form a diverse family, going from very strong logical entailments, through standard IF-THENs, down to probability or to mere suggestion or insinuation. Even apart from relevance logic, conditionals raise many logical issues.

It shouldn't surprise us that central logical principles raise controversies. Even "I see a chair" raises controversies if we push it far enough. But not all alternative views are equally reasonable. I'd contend that, despite controversies, I really do see a chair. And I'd contend that most assumptions about logic that have been held since Aristotle's time are solid.[1]

[1] I do think, however, that in quantified modal logic there's much to be said for *free logic* (§11.4), which is somewhat deviant. For more on deviant logics, see Graham Priest's *An Introduction to Non-Classical Logic*, 2nd ed. (Cambridge: Cambridge University Press, 2008) and J. C. Beall and Bas van Fraassen's *Possibilities and Paradox* (Oxford: Oxford University Press, 2003).

18 Philosophy of Logic

Philosophy of logic deals with issues about logic that are broadly philosophical, especially metaphysical (about reality) or epistemological (about how we know). Here are examples: Are there abstract entities, and does logic presuppose them? Is logic the key to understanding the structure of reality? How do we know logical laws – are they empirical or true by convention? What is truth, and how do different views on truth affect logic? What is the scope of logic?

18.1 Abstract entities

Metaphysics studies the nature of reality. It considers broad views like *materialism* (only the physical is ultimately real), *idealism* (only the mental is ultimately real), and *dualism* (both the physical and the mental are ultimately real). Another issue is whether there are **abstract entities** – entities, roughly, that are neither physical (like apples) nor mental (like feelings); alleged examples include numbers, sets, and properties.

Logic can quickly bring up issues about abstract entities. Take this argument:

This is green.
This is an apple.
∴ Some apple is green.

In discussing this argument, we may talk about abstract entities:

- The *set* of green things; this set seems to be not physical or mental, but rather an abstract entity.
- The *property* of greenness, which can apply either to the color as experienced or to its underlying physical basis; in either case, greenness seems to be not a concrete mental or physical entity, but rather something more abstract that has physical or mental instances.
- The *concept* of greenness (what terms for "green" in various languages mean).
- The *word* "green" and the *sentence* "This is green," which are abstract patterns with written and auditory instances.

- The *proposition* that this is green, which is the truth claim that we assert using "This is green" in English or similar things in other languages. 0369

Platonists, as logicians use the term, are those who straightforwardly accept the existence of such abstract objects. *Nominalists*, in contrast, are unhappy about such entities and want to restrict what exists to concrete physical or mental entities; they try to make sense of logic while rejecting abstract entities. Intermediate views are possible; maybe we should accept abstract entities, not as independently real entities that we discover, but rather as mental creations or fictions. Disputes about such matters go back to ancient and medieval debates about forms and universals, and continue to rage today.

18.2 Metaphysical structures

Does logic give us the key to understand reality's metaphysical structure? Ludwig Wittgenstein, in his *Tractatus Logico-Philosophicus* (1922), argued that it does. He saw the world as the totality of facts. If we state all the facts, we completely describe reality. Facts are about simple objects. An atomic statement pictures a fact by having its elements mirror the simple objects of the world. Language, when completely analyzed, breaks down into such atomic statements. Complex statements are built from atomic ones using logical connectives like "and," "or," and "not." Wittgenstein invented truth tables to show how this works. Some complex statements, like "It's raining or not raining," are true in all cases, regardless of which atomic statements are true; such statements are certain but lack content.

While Wittgenstein thought atomic statements were the simplest truths, he didn't say whether these were about physical facts or experiences. In either case, complex statements are constructible out of atomic statements using the logical connectives of propositional logic (Chapter 6). Statements not so constructible are nonsensical. Wittgenstein thought that most philosophical issues (for example, about values or God) were nonsensical. Paradoxically, he thought his own theory (starting with his claim that the world is the totality of facts) is nonsensical too. He ended on a mystical note: the most important things in life (his own theory, values, God, the meaning of life) cannot be put into words.

Bertrand Russell, while impressed by Wittgenstein's views, tried to make them more sensible and less paradoxical. Russell's *logical atomism* held that an ideal language – one adequate to describe reality completely – must be based on quantificational logic (Chapters 8 and 9) and thus must include quantifiers like "all" and "some." It must also include terms that refer to the ultimately simple elements of reality – which include objects, properties, and relations. He debated whether the basic entities of the world were physical, or mental, or perhaps something neutral between the two.

Russell thought ordinary language can lead us into bad metaphysics (§9.6). Suppose you say "There's nothing in the box." Some might see "nothing" as the name of a mysterious object in the box. This is wrong. Instead, the sentence just 0370 means "It's false that there's something in the box." Or suppose you say "The average American has 2.4 children." While "the average American" doesn't refer to an actual entity, the sentence is meaningful; it asserts that the average number of children that Americans have is 2.4. "Nothing" and "the average American" are **logical constructs**; they're mere ways of speaking and don't directly refer to objects. Russell went on to ask whether sets, numbers, material objects, persons, electrons, and experiences were real entities or logical constructs. Logical analysis is the key to answering such questions. We must see, for example, whether statements about material objects can be reduced to sensations, or whether statements about minds can be analyzed as about behavior.

In a similar spirit, Willard Quine pursued **ontology**, about what kinds of entity ultimately exist. His slogan, "To be is to be the value of a bound variable," tried to clarify ontological disputes. It means that the entities our theory commits us to are those that our quantified variables (like "for all x" and "for some x") must range over for our statements to be true. So if we say, "There's some feature that Shakira and Britney have in common," then we must accept *features* (*properties*) as part of our ontology – unless we can show that we're using an avoidable way of speaking (a "logical construct" in Russell's sense). Quine accepted *sets* in his ontology, because he thought they were needed for math and science; in picking an ontology, he appealed to pragmatic considerations. He rejected properties, concepts, and propositions because he thought they were less clear.

Wittgenstein later supported an *ordinary language* approach and rejected his earlier basing of metaphysics on logic. His *Philosophical Investigations* (1953) saw his earlier work as mistakenly imposing ideas on reality instead of fairly investigating it. His slogan became "Don't think, but look!" Don't say that reality *has* to be such and such, because that's what your preconceptions demand; instead, look and see how it is. He now contended that few concepts had strict analyses. His main example was "game," which has no strict definition. Games typically involve a competition between sides, winning and losing, a combination of skill and luck, and so forth. But none of these *family resemblances* is essential; solitaire drops competition, ring-around-the-rosie drops winning or losing, throwing dice drops skill, and chess drops luck. Any strict analysis of "game" is easily refuted by giving examples of games that violate the analysis. We distort language if we think that all statements must be analyzable into simple concepts that reflect metaphysically simple elements of reality. There's no ideal language that perfectly mirrors reality; instead, there are various language games that humans construct for various purposes. Logic is a language game, invented to help us appraise the correctness of reasoning; we distort logic if we see it as giving us a special key to understand the metaphysical structure of reality.

So we see a range of views about the connection of logic with metaphysics, with Wittgenstein holding different views at different times.[1] 0371

18.3 The basis for logical laws

Let's consider **logical laws** like *modus ponens* and non-contradiction:

- *Modus ponens*: If A then B, A, therefore B.
- Non-contradiction: A and not-A cannot both be true, unless A is taken differently in both instances.

Why are such logical laws correct, and how do we know that they're correct? Thinkers have proposed a range of answers. Here we'll consider five: supernaturalism, psychologism, pragmatism, conventionalism, and realism. (§§17.2–17.4 discussed deviant logics that reject these two laws.)

1. **Supernaturalism** holds that all laws of every sort – whether about physics, morality, math, or logic – depend on God. Radical supernaturalists say that God creates the logical laws or at least makes them true. God could make a world where *modus ponens* and the law of non-contradiction fail; and he could violate the law of non-contradiction – for example, by making "You're reading this sentence" and "You're not reading this sentence" both true. So logical laws are contingent: they could have been false. Moderate supernaturalists, on the other hand, say that logical laws express God's perfect nature. God's perfection require that he be consistent, that his created world follow the laws of logic, and that he desire that we be consistent and logical. Since these aspects of God's nature are necessary, the laws of logic are also necessary. Supernaturalists of both sorts hold that God builds the laws of logic into our minds, so these laws appear to us to be "self-evident" when adequately reflected upon.

Critics object that the laws of logic hold for every possible world, including ones where there's no God; so God cannot provide the basis for these laws. Others say that, since beliefs about logic are more certain than beliefs about God, it's wrong to base logic on God. Still others say that God accepts logical laws because they're inherently valid; logical laws aren't valid just because God chooses to accept them (radical supernaturalism) or because they accord with his nature (moderate supernatualism).[2]

[1] For more on logic and metaphysics, see §3.4 (the logical positivist critique of metaphysics), §9.2 (mind and the substitution of identicals), and §§11.2–11.4 (Aristotelian essentialism).

[2] The parallel view in ethics claims that basic moral principles depend on God's will. See my

2. **Psychologism** holds that logical laws are based on how we think. Logic is part of our biology and natural history. Humans evolved to walk on two feet, have hand-eye coordination, communicate by speech, and think logically; these promote survival and are part of our genetic and biological makeup. Radical psychologism says that logic *describes* how we think; logical laws are psychological laws about thinking. Moderate psychologism, in contrast, sees logic as built into us in a more subtle way; we're so built that at reflective moments we see inconsistency and illogicality as defects – even though at other times our thinking may suffer from such defects. When we reflect on our inconsistencies, we tend to 0372 develop an uncomfortable anxiety that psychologists call "cognitive dissonance"; this is as much a part of our biology and natural history as is thirst. So the laws of logic are built into our instincts.

Critics object that radical psychologism, which claims that logical laws describe our thinking, makes it impossible for us to be illogical or inconsistent. But people often reason invalidly or express inconsistent ideas; so logical laws don't necessarily reflect how we think. Moderate psychologism recognizes this; it sees logical laws as reflecting norms about thinking that are built into us and that we recognize at reflective moments. This approach gives a plausible evolutionary and biological explanation of how logic can be instinctive in us; but it fails if it's taken to explain what makes logical laws true or solidly based. Suppose evolution gave us an instinctive belief in the flatness of the earth; it wouldn't follow that the earth actually *was* flat – or that this belief was so solidly based that we couldn't criticize it. Similarly, the instinctiveness of the laws of logic wouldn't make these logical laws correct or solidly based; maybe our instincts on these matters are right or maybe they're wrong – we'd have to investigate further.

There's also a problem with basing our knowledge of logical laws on evolutionary theory. We need logic to appraise the correctness of scientific theories like evolution; so our knowledge of logic cannot without circularity rest on our knowledge of evolutionary theory. In addition, our knowledge of logic is more solidly based than our knowledge of scientific theories.

3. **Pragmatism** holds that logical laws are based on experience. The broad consensus of humanity is that logic works; when we think things out in a logical and consistent way, we're more apt to find the truth and satisfy our needs. This pragmatic test gives the only firm basis for logic or any other way of thinking.

Critics agree that, yes, logical thinking does work. But logic works because its laws hold of inherent necessity; so logical laws cannot be based on experience. Our experience can show us that something *is* true (for example, that this flower is red); but it cannot show us that something *must* be true (that its opposite is *impossible*). Compare logic to mouse traps. We can test various mouse traps to see how well they work; a given trap might catch a mouse or

Ethics and Religion (New York: Cambridge University Press, 2016).

might not – both are possible. But it's not possible for a logical law to fail – for example, for "If A then B" and "A" to both be true while "B" was false. The necessity of logical laws shows that they cannot be based on experience.

Besides, we cannot know that logic works unless we appeal to observation and reasoning – where the reasoning presupposes logical laws. So the pragmatist defense of logical laws is ultimately circular.

4. **Conventionalism** holds that logical laws are based on verbal conventions. We use logical words like "and," "or," "if-then," and "not" according to rules that can be expressed in basic truth tables (§§6.2–6.6). Given these basic truth tables, we can show *modus ponens* to be valid (since its truth table never gives true premises and a false conclusion); we can similarly show the law of non- contradiction to be true (since its truth table comes out as true in all cases). So we can justify logical laws using conventions about what the logical words mean. 0373 Conventionalism explains why logical laws are necessary; if we deny them, we contradict ourselves, since we violate the meaning of words like "and," "or," "if-then," and "not." It also explains how we can know logical laws in an *a priori* manner, independently of sense experience; logical laws are true by virtue of the meaning of words (§§3.6–3.7), and so we can grasp their truth by becoming clear on what they mean. Conventionalism explains the necessity of logical laws without appealing to controversial beliefs about God, evolution, abstract entities, or our ability to grasp abstract truths. Logic's conventionality also allows for alternative logics that are equally correct but follow different conventions.

Critics raise objections to conventionalism. First, the attempt to prove *modus ponens* using truth tables is circular:

> If the truth table for *modus ponens* never gives true premises and a false conclusion, then *modus ponens* is valid.
> The truth table for *modus ponens* never gives true premises and a false conclusion.
> ∴ *Modus ponens* is valid.
>
> If A then B
> A
> ∴ B

This argument itself uses *modus ponens*; so it's circular, since it assumes from the start that *modus ponens* is valid. Second, conventionalism confuses the logical laws (which are necessary truths) with how we express them (using language conventions). If we changed our language, the logical laws would still be true, but we'd have to express them using different words. Third, conventionalism makes logical laws too arbitrary, since they could fail if we changed our conventions; for example, both *modus ponens* and the law of non-contradiction fail on some many-valued conventions (§17.1). But logical laws seem to have an inherent correctness that doesn't depend on which language conventions we adopt.

5. **Realism** holds that logical laws are objective, independent, abstract truths. We *discover* logical laws; we don't construct or create them. Logical laws aren't reducible to the mental, the physical, usefulness, or conventions. Logical laws govern our world, and every possible world, because violating them is impossible; it cannot be, for example, that A and not-A are both true. Logical laws become self-evident to us when adequately reflected upon. This doesn't mean that logical intuitions are infallible; beginning logic students tend to have poor intuitions about whether an argument is valid. But logical intuitions can be trained; we can test proposed inference forms through concrete examples where the validity or invalidity is more obvious. The best evidence for a logical principle is that a well-trained mind finds it evident and can't find counterexamples.

Critics object that realism makes logical laws too mysterious. Suppose you're a materialist: you hold that all facts are expressible in the language of physics and chemistry. How do objective, irreducible logical facts fit into such a universe? Are logical facts composed of chemicals, or what sort of weird thing are they? And how could we ever know such mysterious logical facts? In addition, objective, abstract logical laws seem to presuppose abstract entities (§18.1), which 0374 have no place in a materialistic world. A dualist view that accepts only mind and matter would have similar doubts about realism.

Logicians for the most part (except for deviant logicians – see Chapter 17) agree on the logical laws. But logicians differ widely on what these laws are based on and how we can know them to be correct.

18.4 Truth and paradoxes

Truth is important to logic. A *valid argument* is often defined as one in which it's impossible to have the premises all true and conclusion false. Truth comes up further in propositional logic (with truth tables and the truth-assignment test) and in refutations of invalid arguments (which are possible situations making the premises all true and conclusion false).

There are many issues about truth. For example, is classical logic right in assuming that statements are true or false, but not both, and that *true* and *false* are the only truth values? Some deviant logics deny these assumptions (Chapter 17).

What do "true" and "false" apply to? Suppose you point to a green apple and say "This is green." Is what is true the *sentence* "This is green," or perhaps the sentence as used on this occasion (where you point to a certain object)? If so, then is this sentence concrete physical marks or sounds, or is it a more abstract pattern that has written or auditory instances? Or perhaps what is true-or-false is not sentences, but rather *propositions*, which are *assertions* that we use language to make. But then are propositions something mental, or are they abstract entities, like the meaning of "This is green"?

What does "true" mean? On different views, being "true" is:

- corresponding to the facts (correspondence theory),
- cohering with our other beliefs (coherence theory),
- being useful to believe (pragmatist theory),
- being verified (verification theory), or
- being what we'd agree to under cognitively ideal conditions (ideal consensus theory); or perhaps
- "It's true that A" is just a wordy way to assert A (redundancy theory).

The pragmatist and verification analyses reject the law of excluded middle, since it can happen that neither a statement nor its negation is useful or verified. These two analyses could also support many-valued logic (§17.1), since a statement can be useful or verified to various degrees. Thus different answers to "What is truth?" can support different logics.

Alfred Tarski proposed an adequacy condition, called "convention T," that any definition of truth must satisfy; here's an example: 0375

The sentence "Snow is white" is *true*, if and only if snow *is* white.

This equivalence raises problems for definitions that water down truth's objectivity. For example, the view that "true" just means "accepted in our culture" leads to an absurdity. Imagine a tropical island where snow *is* white (in high-mountain cracks that are never visited or seen) but yet people don't believe that it's white; on the proposed view, *snow could be white while "Snow is white" wasn't true* – which is absurd. A similar objection works against the pragmatist and verification views. Imagine that "Snow is white" was neither useful to believe nor verified; then, on pragmatism or verificationism, *snow could be white while "Snow is white" wasn't true* – which is absurd.

Further issues are raised by the **liar paradox**, a statement that asserts its own falsity (and so appears to be both true and false). Consider claim P:

(P) P is false.

Is P true? Then things must be as P says they are, and thus P has to be false. Is P false? Then things are as P says they are, and thus P has to be true. So if P is either true or false, then it has to be both true and false.

Graham Priest and others claim that P is *both true and false*, which requires rejecting Aristotle's law of non-contradiction (§17.2). The more common view is that P is *neither true nor false*, which requires rejecting or qualifying Aristotle's law of excluded middle. Bertrand Russell proposed a *theory of types* that outlaws certain forms of self-reference. Very roughly, there are ordinary objects (type 0), properties of these (type 1), properties of these properties (type 2), and so on. Any meaningful statement can talk only about objects of a lower type; so no speech can talk meaningfully about itself. P violates this condition, and so is meaningless – and thus neither true nor false.

But Russell's view seems to refute itself. "Any meaningful statement can talk only about objects of a lower type," to be useful, has to restrict *all statements, of every type*; but then it violates its own rule and declares itself meaningless.

Tarski, to deal with the paradox, proposed that no language can contain its own truth predicate; to ascribe truth or falsity to a statement in a given language, we must ascend to a higher-level language, called the *metalanguage*. P violates this condition and so is meaningless – and thus neither true nor false.

Opponents say Tarski's view is too restrictive. English and other languages *do* contain their own truth predicates, and they need to do this for many purposes. So it would be better to have a less sweeping restriction to take care of the liar paradox. But there's little agreement about what this restriction should be.

Epimenides of Crete in the sixth century BC proposed the liar paradox, and St Paul mentioned it in his letter to Titus (1:12). It has been widely discussed ever since. While most logicians think that a theory of truth must deal with the paradox, how best to do this is still unclear. 0376

18.5 Logic's scope

"Logic" is often defined in ways like "the analysis and appraisal of arguments" or "the study of valid reasoning." The term "logic" can be used in a narrow and a broad sense. *Logic in the narrow sense* is the study of deductive reasoning, which is about what logically follows from what. *Logic in the broad sense* includes also various other studies that relate to the analysis and appraisal of arguments, like informal logic, inductive logic, metalogic, and philosophy of logic (Chapters 3–5, 15, and 18).

Even taking "logic" in this narrow deductive sense, there's still some unclarity on what it includes. Suppose you say, "I have \$30; therefore I have more than \$20." Is this part of logic, part of math, or both?

Willard Quine suggested that we limit "logic" to classical propositional and quantificational logic (Chapters 6 to 9), which he saw as fairly uncontroversial and as focusing on topic-neutral terms like "and" and "not" that arise in every area of study. Modal and deontic logic (Chapters 10 to 12) focus on terms like "necessary" and "ought" that are too colorful and topic-specific to be part of logic; these areas, if legitimate at all (and he had doubts) are part of philosophy in general, not part of logic. Mathematical extensions, like set theory and axiomatizations of arithmetic, belong to math. And deviant logics (Chapter 17) are illegitimate.

Most logicians today tend to use "(deductive) logic" in a broader way that's hard to pin down. Deductive logic is commonly taken to include, besides syllogisms and classical symbolic logic, extensions like modal and deontic logic, deviant logics, and sometimes even mathematical extensions like set theory. Logic is part of at least three disciplines – philosophy, math, and computer science – which approach it from different angles. Any attempt to give sharp and final boundaries to the term "logic" would be artificial.[1]

[1] For more on philosophy of logic, see Willard Quine's *Philosophy of Logic*, 2nd ed. (Cambridge, Mass.: Harvard University Press, 1986), which is a good introduction from an influential and controversial thinker, and Colin McGinn's *Logical Properties: Identity, Existence, Predication, Necessity, Truth* (Oxford: Clarendon, 2000), which gives an opposing view.

For Further Reading

If you've mastered this book and want more, consult my *Historical Dictionary of Logic* (Lanham, Md.: Scarecrow Press, 2006). This brief encyclopedia of logic has nontechnical, alphabetized articles on branches of logic, figures and historical periods, specialized vocabulary, controversies, and relationships to other disciplines – and a 13-page chronology of major events in the history of logic. It also has a 52-page bibliography of readings in logic, a list of 63 recommended works in various categories, and this smaller list of very helpful works:

- P. H. Nidditch's *The Development of Mathematical Logic* (London: Routledge & Kegan Paul, 1962): the history of logic from Aristotle onward.
- Willard Quine's *Philosophy of Logic*, 2nd ed. (Cambridge, Mass.: Harvard University Press, 1986): a contentious introduction from a major thinker.
- Colin McGinn's *Logical Properties: Identity, Existence, Predication, Necessity, Truth* (Oxford: Clarendon, 2000): an opposing view from Quine's.
- Graham Priest's *An Introduction to Non-Classical Logic*, 2nd ed. (Cambridge: Cambridge University Press, 2008) a defense of deviant logic by its most eloquent defender (technical parts may be skipped).
- Ian Hacking's *An Introduction to Probability and Inductive Logic* (Cambridge: Cambridge University Press, 2001): a solid introduction.
- George Boolos and Richard Jeffrey's *Computability and Logic*, 3rd ed. (Cambridge: Cambridge University Press, 1989): topics like Turing machines, uncomputable functions, the Skolem-Löwenheim theorem, and Gödel's theorem – technical but clear and doesn't assume much math.

If you're just starting, you might pick one or two of these that interest you. For further suggestions, consult my *Historical Dictionary of Logic*.

As advanced students go through various chapters, they might want to pursue further readings *in this book*. Chapter 6 (basic propositional logic) goes well with metalogic §§15.1–2, deviant logic Chapter 17, and philosophy of logic §18.4. Chapter 7 (propositional proofs) goes well with metalogic §§15.3–5 and perhaps informal and inductive Chapters 3 to 5. Chapters 8 and 9 (quantificational logic) go well with metalogic §15.6, history of logic Chapter 16, philosophy of logic §§18.1–3, and syllogisms Chapter 2. And Chapters 10 to 14 (modal/deontic/belief logic and a formalized ethical theory) go well with history of logic §16.5 and philosophy of logic §18.5.

Answers to Selected Problems

For each exercise set in the book, answers are given for problems 1, 3, 5, 10, 15, 20, 25, and so on. The teachers manual (see Preface) has answers to the other problems.

Chapter 2 answers

2.1a

1. t is S
3. no L is B
5. all D is H
10. a is s
15. m is A

2.2a

1. This isn't a syllogism, because "D" and "E" occur only once.
3. This isn't a syllogism, because "Y" occurs three times and "G" occurs only once.
5. This isn't a syllogism, because "Z is N" isn't a wff.

2.2b

1. w is not **s**
3. no **R** is **S**
5. all **P** is B

2.2c

1. no P* is B*   Invalid
some C is not B*
∴ some C* is P*

3. no H* is B*   Invalid
no H* is D*
∴ some B* is not D

5. ∴ g* is g*   Valid

10. all D* is A   Invalid
∴ all A is D*

2.3a

1. all S* is D   Valid
all D* is U
∴ all S is U*

3. all T* is C   Valid
no C* is R*
∴ no T is R

5. all M* is R   Valid
some P is M
∴ some P* is R*

10. all S* is Y   Invalid
m is Y
∴ m* is S*

15. all N* is L   Valid
m is N
∴ m* is L*

20. b is W   Invalid
u is W
∴ u* is b*

25. some S is W   Valid
all S* is L
all L* is H
∴ some W* is H*

2.3b

1. We can't prove either "Bob stole money" or "Bob didn't steal money." 2 & 6 yield no valid argument with either conclusion.
3. 4 & 8 & 9 prove David stole money: "d is W, all W is H, all H is S ∴ d is S."
5. This would show that our data was inconsistent and so contains false information. 0379

2.4a

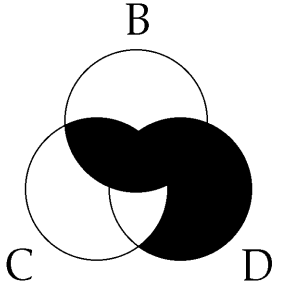
1. all J is F
3. all S is R
5. some H is L
10. no S is H
15. all M is B
20. some H is not G

2.5a

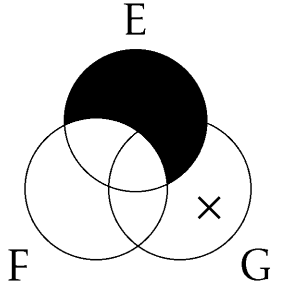
1. "No human acts are free" or "No free acts are human acts."
3. "Some free acts are determined" or "Some determined acts are free."
5. No conclusion validly follows.
10. "No culturally taught racial feelings are rational" or "No rational thing is a culturally taught racial feeling."
15. "Some who like raw steaks like champagne" or "Some who like champagne like raw steaks."
20. "No basic moral norms are principles based on human nature" or "No principles based on human nature are basic moral norms."
25. "No moral judgments are objective truths" or "No objective truths are moral judgments."

2.6a

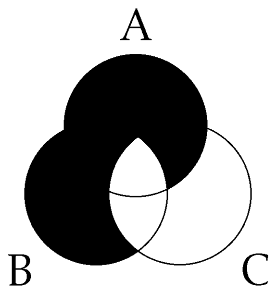
1. no B is C   Valid
all D is C
∴ no D is B



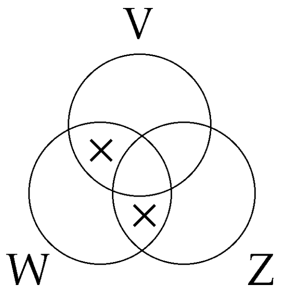
3. all E is F   Valid
some G is not F
∴ some G is not E



5. all A is B   Valid
all B is C
∴ all A is C



10. some V is W   Invalid
some W is Z
∴ some V is Z



2.7a

1. all R* is G   Valid
all G* is T
all T* is V
all V* is U
∴ all R is U*

3. g is A   Valid
all A* is R
no R* is C*
∴ g* is not C

5. no S* is A*   Valid
all W* is A
∴ no S is W

Premise 2 (implicit but false) is "All garments that should be worn next to the skin while skiing are garments that absorb moisture."

10. all P* is O   Valid
all O* is E
no M* is E*
∴ no M is P

15. e is C   Invalid
all S* is C
∴ e* is S*

20. all N* is C   Valid
no E* is C*
g is E
∴ g* is not N

Premise 3 (implicit) is "'God exists' is an existence claim."

25. all D* is F   Valid
some P is not F*
∴ some P* is not D

Chapter 3 answers

3.1a

1. "Cop" is negative. "Police" is more neutral.
3. "Heroic" is positive. These are negative: "reckless," "foolhardy," "brash," "rash," "careless," "imprudent," and "daredevil."
5. "Elderly gentleman" is positive. "Old man" is negative. 0380
10. "Do-gooder" is negative. "Person concerned for others" and "caring human being" are positive.
15. "Booze" is negative or neutral. "Cocktail" is positive, while "alcohol," "liquor," and "intoxicant" are neutral.
20. "Babbling" is negative. "Talking," "speaking," and "discussing" are neutral.
25. "Bribe" is negative. "Payment" and "gift" are neutral or positive.
30. "Whore" is negative. "Prostitute" is more neutral.

3.2a

1. A false statement that you think is true isn't a lie.
3. (1) One who believes in God may not make God his or her ultimate concern. (2) One may have an ultimate concern (such as making money) without believing in God. (3) "Object of ultimate concern" is relative in a way that "God" isn't: "Is there an object of ultimate concern?" invites the question "For whom?" – while "Is there a God?" doesn't.
5. Since "of positive value" is no more clearly understood than "good," this definition does little to clarify what "good" means. And there's the danger of circularity if we go on to define "of positive value" in terms of "good."
10. (1) If I believe that Michigan will beat Ohio State next year, it still might not be true. (2) If "true" means "believed," then both these statements are true (since both are believed by someone): "Michigan will beat Ohio State next year" and "Michigan won't beat Ohio State next year." (3) "Believed" is relative in a way that "true" isn't: "Is this believed?" invites the question "By whom?" – while "Is this

true?" doesn't.
15. This set of definitions is circular.

3.2b

1. This is true according to cultural relativism. Sociological data can verify what is "socially approved," and this is the same as what is "good."
3. This is true. The norms set up by my society determine what is good in my society, so these norms couldn't be mistaken.
5. This is undecided. If our society approves of respecting the values of other societies, then this respect is good. But if our society disapproves of respecting the values of other societies, then this respect is bad.
10. This is true according to CR.
15. This is false (and self-contradictory) according to cultural relativism.
20. This is undecided, since cultural relativism leaves unspecified which of these various groups is "the society in question."

3.4a

1. This is meaningful on LP (it could be verified) and PR (it could make a practical difference in terms of sensations or choices).
3. This is meaningful on both views.
5. This is probably meaningless on both views (unless the statement is given some special sense).
10. This is meaningless on LP (at least on the version that requires public verifiability). It's meaningful on PR (since its truth could make a practical difference to Manuel's experience).
15. Since this (LP) isn't able to be tested empirically, it's meaningless on LP. [To avoid this result, a positivist could claim that LP is true by definition and hence analytic (§3.6). Recall that LP is qualified so that it applies only to synthetic statements. But then the positivist has to use "meaningless" in the unusual sense of "synthetic but not empirical" instead of in the intended sense of "true or false." This shift takes the bite out of the claim that a statement is "meaningless." A believer can readily agree that "There is a God" is "meaningless" if all this means is that "There is a God" isn't synthetic-but-not-empirical.] It's meaningful on PR (its truth could make a difference to our choices about what we ought to believe).

3.5a

(These answers were adapted from those given by my students)

1. "Is ethics a **<u>science</u>**?" could mean any of the following:
- Are ethical judgments true or false independently of human feelings and opinions? Can the truth of some ethical judgments be known? 0381
- Can ethics be systematized into a set of rules that will tell us unambiguously what we ought to do in all (or most) cases?
- Can ethical principles be proved using the methods of empirical science?
- Is there some rational method for arriving at ethical judgments that would

lead people to agree on their ethical judgments?
- Can a system of ethical principles be drawn up in an axiomatic form, so that ethical theorems can be deduced from axioms accessible to human reason?

3. "Is this belief part of **common sense**?" could mean any of the following:
- Is this belief accepted instinctively or intuitively, as opposed to being the product of reasoning or education?
- Is this belief so entrenched that subtle reasoning to the contrary, even if it seems flawless, has no power to convince us?
- Is this belief something that people of good "horse sense" will accept regardless of their education?
- Is this belief obviously true?
- Is this belief universally accepted?

[In each case we could further specify the group we are talking about – for example, "Is this belief obviously true to anyone who has ever lived (to all those of our own country, or to practically all those of our own country who haven't been exposed to subtle reasoning on this topic)?"]

5. "Are values **relative** (or **absolute**)?" could mean any of the following:
- Do different individuals and societies disagree (and to what extent) on values?
- Do people disagree on basic moral principles (and not just on applications)?
- Are all (or some) values incapable of being proved or rationally argued?
- Is it wrong to claim that a moral judgment is correct or incorrect rather than claiming that it's correct or incorrect relative to such and such a group? Do moral judgments express social conventions rather than truths that hold independently of such conventions?
- Do right and wrong always depend on circumstances (so that no sort of action could be always right or always wrong)?
- In making concrete moral judgments, do different values have to be weighed against each other?
- Are all things that are valued only valued as a means to something else (so that nothing is valued for its own sake)?

10. "Is that judgment based on **reason**?" could be asking whether the judgment is based on the following:
- Self-evident truths, the analysis of concepts, and logical deductions from these (reason versus experience).
- The foregoing plus sense experience, introspection, and inductive arguments (reason versus faith).
- Some sort of thinking or experience or faith (as opposed to being based on mere emotion).
- The thinking and experience and feelings of a sane person (as opposed to those of an insane person).
- An adequate and impartial examination of the available data.
- A process for arriving at truth in which everyone correctly following it would arrive at the same conclusions.
- What is reasonable to believe, or what one ought to believe (or what is per-

missible to believe) from the standpoint of the seeking of truth.

[We could be asking whether a given person bases his or her judgment on one of the foregoing, or whether the judgment in question could be based on one of the foregoing.]

15. "Do you have a **soul**?" could mean any of the following:

- Do you have a personal identity that could in principle survive death and the disintegration of your body?
- Are you capable of conscious thinking and doing?
- Would an exhaustive description of your material constituents and observable behavior patterns fail to capture important elements of what you are?
- Are you composed of two quite distinct beings – a thinking being without spatial dimensions and a material being incapable of thought?
- Are you capable of caring deeply about anything?
- Are you still alive?

3.6a

1. Analytic.
3. Synthetic.
5. Analytic.
10. Analytic.
15. Analytic.
20. Most philosophers think this is synthetic. St Anselm, Descartes, and Charles 0382 Hartshorne argued that it was analytic. See examples 3 and 4 of §6.7b, and examples 9 and 26 of §10.3b.
25. Most say synthetic, but some say analytic.

3.7a

1. *A priori*.
3. *A posteriori*.
5. *A priori*.
10. *A priori*.
15. *A priori*.
20. Most philosophers think this could only be known *a posteriori*. Some philosophers think it can be known *a priori* (see comments on problem 20 of the last section).
25. Most philosophers think this could only be known *a priori*, but a few think it could be known *a posteriori*.

Chapter 4 answers

4.2a

1. Complex question (like "Are you still beating your wife?").

3. Pro–con. The candidate might be a crook. Or an opposing candidate might be even more intelligent and experienced.
5. Appeal to the crowd.
10. Genetic.
15. Appeal to authority.
20. None of the labels fit exactly. This vague claim (what is a "discriminating backpacker"?) is probably false (discriminating backpackers tend to vary in their preferences). The closest labels are "appeal to authority," "appeal to the crowd," "false stereotype," or perhaps "appeal to emotion." There's some "snob appeal" here too, but this isn't one of our categories.
25. *Post hoc ergo propter hoc*.
30. Appeal to opposition.
35. Appeal to emotion.
40. *Post hoc ergo propter hoc*.
45. *Ad hominem* or false stereotype.
50. *Post hoc ergo propter hoc*. The conclusion might still be true, but we'd need a longer argument to show this; many argue, for example, that Bush's deregulation of banking caused the financial crisis.
55. Ambiguous.
60. Black and white, or complex question.

4.2b

1. Complex question.
3. Ambiguity.
5. False stereotype.
10. Appeal to authority.
15. Pro–con.
20. Genetic.
25. Black and white.
30. *Ad hominem*.
35. Appeal to the crowd.
40. Part–whole.
45. Appeal to authority, *ad hominem*, or appeal to emotion.
50. Circular.
55. Complex question.
60. Circular (but it still might be true).

4.3a

(The answers for 3 and 5 are representative correct answers; other answers may be correct.)

1. There are no universal duties.
If everyone ought to respect the dignity of others, then there are universal duties.
∴ Not everyone ought to respect the dignity of others.

3. If we have ethical knowledge, then either ethical truths are provable or there are self-evident ethical truths.

We have ethical knowledge.
Ethical truths aren't provable.
∴ There are self-evident ethical truths.

5. All human concepts derive from sense experience.
The concept of logical validity is a human concept.
∴ The concept of logical validity derives from sense experience.

10. If every rule has an exception, then there's an exception to this idea too; but then some rule doesn't have an exception. Statement 10 implies its own falsity and hence is self-refuting.

15. If it's impossible to express truth in human concepts, then statement 15 is false. Statement 15 implies its own falsity and hence is self-refuting.

4.4a

(These are examples of answers and aren't the only "right answers.")

1. If the agent will probably get caught, then offering the bribe probably isn't in the agent's self-interest.
The agent will probably get caught. (One might give inductive reasoning for this.)
0383
∴ Offering the bribe probably isn't in the agent's self-interest.

3. Some acts that grossly violate the rights of some maximize good consequences (in the sense of maximizing the total of everyone's interests).
No acts that grossly violate the rights of some are right.
∴ Some acts that maximize good consequences aren't right.

5. Any act that involves lying is a dishonest act (from the definition of "dishonest").
Offering the bribe involves lying (falsifying records, and the like).
∴ Offering the bribe is a dishonest act.

10. Science adequately explains our experience.
If science adequately explains our experience, then the belief that there is a God is unnecessary to explain our experience.
∴ The belief that there is a God is unnecessary to explain our experience.

Or: Science doesn't adequately explain certain items of our experience (why these scientific laws govern our universe and not others, why our universe exhibits order, why there exists a world of contingent beings at all, moral obligations, and so on).
If science doesn't adequately explain certain items of our experience, then the belief that there is a God is necessary to explain our experience.
∴ The belief that there is a God is necessary to explain our experience.

15. The idea of logical validity is an idea gained in our earthly existence.

The idea of logical validity isn't derived from sense experience.
∴ Some ideas gained in our earthly existence don't derive from sense experience.

Chapter 5 answers

5.2a

1. There are 32 such cards out of the 103 remaining cards. So your probability is 32/103 (about 31.1 percent).
3. Coins have no memory. The probability of heads is 50 percent.
5. The probability that Michigan will win the Rose Bowl is 80 percent times 60 percent times 30 percent, or 14.4 percent.
10. You get a number divisible by three 12 out of 36 times. You don't get it 24 out of 36 times. Thus, mathematically fair betting odds are 2 to 1 (24 to 12) against getting a number divisible by three.
15. In 100 such cases, Ohio State would pass 60 times and run 40 times. If we set up to stop the pass, we'd stop them 58 times out of 100 [(60 • 70 percent) + (40 • 40 percent)]. If we set up to stop the run, we'd stop them 62 times out of 100 [(60 • 50 percent) + (40 • 80 percent)]. So we should set up to stop the run.

5.3a

1. You shouldn't believe it. It's only 12.5 percent (50 • 50 • 50 percent) probable.
3. You shouldn't believe it. It's 37.5 percent probable, since it happens in 3 of the 8 possible combinations.
5. You shouldn't believe it. It's not more probable than not; it's only 50 percent probable.
10. You should buy the Enormity Incorporated model. If you buy the Cut-Rate model, there's an expected replacement cost of $360 ($600 times 60 percent) in addition to the $600 purchase price. This makes the total expected cost $960. The expected cost on the Enormity Incorporated model is $900.

5.4a

1. This is a poor argument, since the sample has little variety.
3. This is a poor argument, since the sample is very small and lacks variety.
5. This is a good inductive argument (if you aren't in the polar regions where the sun doesn't come up at all for several weeks in the winter). In standard form, the argument goes: "All examined days are days when the sun comes up; a large and varied group of days has been examined; tomorrow is a day; so probably tomorrow is a day when the sun comes up."
10. This weakens the argument. Some students cram logic mainly for the Law School Admissions Test (since this test contains many logic problems). You might not have known this, however.

5.5a

1. This doesn't affect the strength of the argument, since the color of the book has little to do with the contents. 0384
3. This weakens the argument. It's less likely that a course taught by a member of the math department would include a discussion of analogical reasoning.
5. This weakens the argument. An abstract approach that stresses theory is less likely to discuss analogical reasoning.
10. This weakens the argument. A book with only 10 pages on inductive reasoning is less likely to include analogical reasoning.
15. This weakens the argument, since it's a significant point of difference between the two cases.

5.7a

1. Using the method of agreement, we conclude that either having a few drinks causes a longer reaction time, or having a longer reaction time causes a person to have a few drinks. The second alternative is less likely in terms of our background information. So we conclude that having a few drinks probably causes a longer reaction time.
3. The method of agreement seems to lead to the conclusion that the soda caused the hangover. However, we know that scotch, gin, and rum all contain alcohol. So soda isn't the only factor common to all four cases; there's also the alcohol. So the method of agreement doesn't apply here. To decide whether the soda or the alcohol caused the hangover, Michelle would have to experiment with drinking soda but no alcohol, and drinking alcohol but no soda.
5. Using the method of agreement, we'd conclude that either factor K caused cancer or cancer caused factor K. If we found some drug to eliminate factor K, then we could try it and see whether it eliminates cancer. If eliminating factor K eliminated cancer, then it's likely that factor K caused cancer. But if factor K came back after we eliminated it, then it's likely that cancer caused factor K.
10. Using the method of disagreement, we'd conclude that eating raw garlic doesn't by itself necessarily cause mosquitoes to stop biting you.
15. Using the method of agreement, we'd conclude that either the combination of factors (heating or striking dry matches in the presence of oxygen) causes the match to light, or else the lighting of the match causes the combination of factors. The latter is implausible (it involves a present fire causing a past heating or striking). So probably the combination of factors causes the match to light.
20. By the method of variation, it's likely that an increase in the electrical voltage is the cause of the increase in the electrical current, or the electrical current is the cause of the electrical voltage, or something else caused them both. We know (but perhaps little Will doesn't) that we can have a voltage without a current (such as when nothing is plugged in to our electrical socket) but we can't have a current without a voltage. So we'd think that voltage causes current (and not vice versa) and reject the "electrical current is the cause of the electrical voltage" alternative. So we'd conclude that probably an increase in the electrical voltage is the cause of the increase in the electrical current, or else some other factor (Will's curiosity, for example) caused both increases.
25. By the method of difference, wearing a single pair of socks probably is (or is

part of) the cause of the blisters, or the blisters are (or are part of) the cause of wearing a single pair of socks. The latter is impossible, since a present event can't cause a past event. So probably wearing a single pair of socks is (or is part of) the cause of the blisters. Since we know that we don't get blisters from wearing a single pair of socks without walking, we'd conclude that wearing a single pair of socks is only part of the cause of the blisters.

5.8a

1. The problem is how to do the experiment so that differences in air resistance won't get in the way. We could build a 100-foot tower on the moon (or some planet without air), drop a feather and a rock from the top, and see if both strike the ground at the same time. Or we might go to the top of a high building and drop rocks of different weights to see if they land at about the same time (with perhaps very minor time differences dues to minor differences in air resistance between rocks).
3. We could study land patterns (hills, rock piles, eccentric boulders, and so on) left by present-day glaciers in places like Alaska, compare land patterns of areas that we are fairly sure weren't covered by glaciers, and compare both with those of Wisconsin. Mill's method of agreement might lead us 0385 to conclude that glaciers probably caused the land patterns in Wisconsin. To date the glacier, we'd have to find some "natural calendar" (such as the yearly rings in tree trunks, yearly sediment layers on the bottoms of lakes, corresponding layers in sedimentary rocks, or carbon breakdown) and connect it with Wisconsin climatic changes or land patterns.
5. We could give both groups an intelligence test. The problem is that the first child might test higher, not because of greater innate intelligence, but because of differences in how the first and the last child are brought up. (The last child, but not the first, is normally brought up with other children around and by older parents.) To eliminate this factor, we might test adopted children. If we find that a child born first and one born last tend to test equally (or unequally) in the same sort of adoptive environment, then we could conclude that the two groups tend (or don't tend) to have the same innate intelligence.
10. See the answer to problem 3. Any data making statement 3 probable would make 10 improbable. In addition, if we found any "natural calendar" that gives a strong inductive argument concerning any events occurring over 5,000 years ago, this also would make 10 unlikely. [Of course, these are only inductive arguments; it's possible for the premises to be all true and conclusion false.]

Chapter 6 answers

6.1a

1. $\sim(A \bullet B)$
3. $((A \bullet B) \vee C)$
5. $((A \supset B) \vee C)$
10. $(A \supset \sim(\sim B \bullet \sim C))$

15. (~(E ∨ P) ⊃ ~R)
20. E ["(M ∨ F)" is wrong, since the English sentence doesn't mean "Everyone is male or everyone is female."]

6.2a

1. 1
3. 1
5. 0
10. 1
15. 0

6.3a

1. ~(1 • 0) = ~0 = 1
3. ~(~1 • ~0) = ~(0 • 1) = ~0 = 1
5. (~0 ≡ 0) = (1 ≡ 0) = 0
10. (~1 ∨ ~(0 ⊃ 0)) = (0 ∨ ~1) = (0 ∨ 0) = 0
15. ~((1 ⊃ 1) ⊃ (1 ⊃ 0)) = ~(1 ⊃ 0) = ~0 = 1

6.4a

1. (? • 0) = 0
3. (? ∨ ~0) = (? ∨ 1) = 1
5. (0 ⊃ ?) = 1
10. (? ⊃ ~0) = (? ⊃ 1) = 1

6.5a

1.

P Q	(P ≡ ~Q)
0 0	0
0 1	1
1 0	1
1 1	0

3.

P Q R	(P ∨ (Q • ~R))
0 0 0	0
0 0 1	0
0 1 0	1
0 1 1	0
1 0 0	1
1 0 1	1
1 1 0	1
1 1 1	1

5.

P Q	((P ≡ Q) ⊃ Q)

0 0	0
0 1	1
1 0	1
1 1	1

6.6a

1. Invalid: second row has 110.

C D	(C ⊃ D),	D	∴	C
0 0	1	0		0
0 1	**1**	**1**		**0**
1 0	0	0		1
1 1	1	1		1

3. Valid: no row has 110.

T B	(T ⊃ B),	(T ⊃ ~B)	∴	~T
0 0	1	1		1
0 1	1	1		1
1 0	0	1		0
1 1	1	0		0

5. Invalid: row 4 has 1110. (I once got a group together but couldn't get Grand Canyon backcountry reservations. So we instead explored canyons near Escalante, Utah. This made R = 0, T = 1, and E = 1.) 0386

R T E	((R • T) ⊃ E),	T	~R	∴	~E
0 0 0	1	0	1		1
0 0 1	1	0	1		0
0 1 0	1	1	1		1
0 1 1	**1**	**1**	**1**		**0**
1 0 0	1	0	0		1
1 0 1	1	0	0		0
1 1 0	0	1	0		1
1 1 1	1	1	0		0

10. Invalid: row 1 has 110.

S E R	(S ⊃ (E • ~R)),	~E	∴	R
0 0 0	**1**	**1**		**0**
0 0 1	1	1		1
0 1 0	1	0		0
0 1 1	1	0		1
1 0 0	0	1		0
1 0 1	0	1		1
1 1 0	1	0		0
1 1 1	0	0		1

6.7a

1. $\sim(N^1 \equiv H^1) \neq 1$ Valid
$N^1 = 1$
$\therefore \sim H^1 = 0$

3. $((T \vee M^1) \supset Q^0) \neq 1$ Valid

$M^1 = 1$
$\therefore Q^0 = 0$

5. $((L^0 \bullet F^1) \supset S^1) = 1$ Invalid
$S^1 = 1$
$F^1 = 1$
$\therefore L^0 = 0$

10. $(\sim T^0 \supset (P^1 \supset J^0)) \neq 1$ Valid
$P^1 = 1$
$\sim J^0 = 1$
$\therefore T^0 = 0$

15. $A^1 = 1$ Valid
$\sim A^1 \neq 1$
$\therefore B^0 = 0$

(An argument with inconsistent premises is always valid: if the premises can't all be true, we can't have premises all true and conclusion false. But such an argument can't be sound, since the premises can't all be true. This argument is controversial – see §17.2.)

6.7b

1. C   Valid
A
$((C \bullet A) \supset (F \vee I))$
$\sim I$
$\therefore F$

3. $((U \bullet \sim R) \supset C)$ Valid
$\sim C$
U
$\therefore R$

5. $((S \bullet \sim M) \supset D)$ Valid
$\sim D$
S
$\therefore M$

10. $(I \supset (U \vee \sim P))$ Invalid
$\sim U$
$\sim P$
$\therefore I$

15. $((M \bullet S) \supset G)$ Valid
S
$\sim G$

∴ ~M

20. ((I • ~D) ⊃ R)   Valid
~D
I
∴ R

6.8a

1. (S ⊃ (Y • I))
3. (Q ∨ R)
5. (~T ⊃ ~P)
10. (A ⊃ E) *or, equivalently,* (~E ⊃ ~A)
15. (S ⊃ W)

6.9a

1. (S ⊃ ~K)   Valid
K
∴ ~S

The implicit premise 2 is "We can know something that we aren't presently sensing."

3. ((B • ~Q) ⊃ O)   Invalid
~B
∴ ~O

5. (S ⊃ A)   Valid
~A
∴ ~S

The implicit premise 2 is "The basic principles of ethics aren't largely agreed upon by intelligent people who have studied ethics."

10. (K ⊃ (P ∨ S))   Valid
~P
~S
∴ ~K

15. (O ⊃ (H ∨ C))   Valid
~C
O
∴ H 0387

20. G   Valid
~S
((M • G) ⊃ S)

∴ ~M

6.10a

1. P, U
3. no conclusion
5. ~R, ~S
10. H, I
15. no conclusion
20. no conclusion

6.11a

1. ~T
3. ~B
5. no conclusion
10. no conclusion
15. F
20. Y

6.12a

1. ~U
3. no conclusion
5. P, ~Q
10. ~A
15. no conclusion

Chapter 7 answers

7.1a

1. Valid

```
* 1   (A ⊃ B)
  [ ∴ (~B ⊃ ~A)
* 2 ┌ asm: ~(~B ⊃ ~A)
  3 │ ∴ ~B    {from 2}
  4 │ ∴ A    {from 2}
  5 └ ∴ ~A    {from 1 and 3}
  6 ∴ (~B ⊃ ~A)    {from 2; 4 contradicts 5}
```

3. Valid

```
* 1   (A ⊃ B)
* 2   (~A ⊃ B)
  [ ∴ B
  3 ┌ asm: ~B
  4 │ ∴ ~A    {from 1 and 3}
```

5 └ ∴ A {from 2 and 3}
6 ∴ B {from 3; 4 contradicts 5}

5. Valid

* 1 (A ∨ B)
* 2 (A ⊃ C)
* 3 (B ⊃ D)
[∴ (C ∨ D)
* 4 ┌ asm: ~(C ∨ D)
5 │ ∴ ~C {from 4}
6 │ ∴ ~D {from 4}
7 │ ∴ ~A {from 2 and 5}
8 │ ∴ B {from 1 and 7}
9 └ ∴ ~B {from 3 and 6}
10 ∴ (C ∨ D) {from 4; 8 contradicts 9}

10. Valid

* 1 (A ⊃ (B ⊃ C))
[∴ ((A • B) ⊃ C)
* 2 ┌ asm: ~((A • B) ⊃ C)
* 3 │ ∴ (A • B) {from 2}
4 │ ∴ ~C {from 2}
5 │ ∴ A {from 3}
6 │ ∴ B {from 3}
* 7 │ ∴ (B ⊃ C) {from 1 and 5}
8 └ ∴ ~B {from 4 and 7}
9 ∴ ((A • B) ⊃ C) {from 2; 6 contradicts 8}

7.1b

1. Valid

* 1 ((H • T) ⊃ B)
2 H
[∴ (T ⊃ B)
* 3 ┌ asm: ~(T ⊃ B)
4 │ ∴ T {from 3}
5 │ ∴ ~B {from 3}
* 6 │ ∴ ~(H • T) {from 1 and 5}
7 └ ∴ ~T {from 2 and 6}
8 ∴ (T ⊃ B) {from 3; 4 contradicts 7}

3. Valid

* 1 (W ⊃ (F ∨ M))
2 ~F
* 3 (M ⊃ A)
[∴ (W ⊃ A)
* 4 ┌ asm: ~(W ⊃ A)
5 │ ∴ W {from 4}
6 │ ∴ ~A {from 4}
* 7 │ ∴ (F ∨ M) {from 1 and 5}
8 │ ∴ ~M {from 3 and 6}

9 └ ∴ M {from 2 and 7}
10 ∴ (W ⊃ A) {from 4; 8 contradicts 9}

5. Valid

* 1 (∼W ⊃ ∼G)
* 2 (∼A ⊃ ∼P)
* 3 (∼W ∨ ∼A)
[∴ (∼P ∨ ∼G)
* 4 ┌ asm: ∼(∼P ∨ ∼G)
5 │ ∴ P {from 4}
6 │ ∴ G {from 4}
7 │ ∴ W {from 1 and 6}
8 │ ∴ A {from 2 and 5}
9 └ ∴ ∼A {from 3 and 7}
10 ∴ (∼P ∨ ∼G) {from 4; 8 contradicts 9}

(This also could be translated without the NOTs – by letting "W," for example, stand for "God doesn't want to prevent evil.")

10. Valid

* 1 (I ⊃ ∼R) 0388
* 2 (∼R ⊃ ∼T)
3 T
* 4 (∼I ⊃ F)
[∴ F
5 ┌ asm: ∼F
6 │ ∴ R {from 2 and 3}
7 │ ∴ ∼I {from 1 and 6}
8 └ ∴ I {from 4 and 5}
9 ∴ F {from 5; 7 contradicts 8}

7.2a

1. Invalid

* 1 (A ∨ B)
[∴ A
2 asm: ∼A
3 ∴ B {from 1 and 2}

∼A, B

3. Invalid

1 ∼(A • ∼B)
[∴ ∼(B • ∼A)
* 2 asm: (B • ∼A)
3 ∴ B {from 2}
4 ∴ ∼A {from 2}

B, ∼A

5. Invalid

```
  1  ((A ⊃ B) ⊃ (C ⊃ D))
* 2  (B ⊃ D)
* 3  (A ⊃ C)
  [ ∴ (A ⊃ D)
* 4  asm: ~(A ⊃ D)
  5  ∴ A    {from 4}
  6  ∴ ~D    {from 4}
  7  ∴ ~B    {from 2 and 6}
  8  ∴ C    {from 3 and 5}
```

A, ~D, ~B, C

10. Invalid

```
* 1  ~(~A • ~B)
  2  ~C
* 3  (D ∨ ~A)
* 4  ((C • ~E) ⊃ ~B)
  5  ~D
  [ ∴ ~E
  6  asm: E
  7  ∴ ~A    {from 3 and 5}
  8  ∴ B    {from 1 and 7}
  9  ∴ ~(C • ~E)    {from 4 and 8}
```

~C, ~D, E, ~A, B

7.2b

1. Invalid

```
  1  (S ⊃ K)
* 2  (M ⊃ K)
  3  M
  [ ∴ S
  4  asm: ~S
  5  ∴ K    {from 2 and 3}
```

M, ~S, K

3. Valid

```
  1  E
* 2  (E ⊃ ~R)
* 3  ((I • ~A) ⊃ R)
  4  I
  [ ∴ A
  5  ┌ asm: ~A
  6  │ ∴ ~R    {from 1 and 2}
* 7  │ ∴ ~(I • ~A)    {from 3 and 6}
  8  └ ∴ A    {from 4 and 7}
  9 ∴ A    {from 5; 5 contradicts 8}
```

5. Valid

* 1 (E ⊃ W)
* 2 (W ⊃ (~R • I))
[∴ (R ⊃ ~E)
* 3 ┌ asm: ~(R ⊃ ~E)
4 │ ∴ R {from 3}
5 │ ∴ E {from 3}
6 │ ∴ W {from 1 and 5}
7 │ ∴ (~R • I) {from 2 and 6}
8 └ ∴ ~R {from 7}
9 ∴ (R ⊃ ~E) {from 3; 4 contradicts 8}

10. Valid

* 1 (G ∨ H)
* 2 ((G • K) ⊃ F)
* 3 (H ⊃ C)
* 4 ((C • K) ⊃ U)
[∴ (K ⊃ (U ∨ F))
* 5 ┌ asm: ~(K ⊃ (U ∨ F))
6 │ ∴ K {from 5}
* 7 │ ∴ ~(U ∨ F) {from 5}
8 │ ∴ ~U {from 7}
9 │ ∴ ~F {from 7}
* 10 │ ∴ ~(G • K) {from 2 and 9}
* 11 │ ∴ ~(C • K) {from 4 and 8}
12 │ ∴ ~G {from 6 and 10}
13 │ ∴ H {from 1 and 12}
14 │ ∴ C {from 3 and 13}
15 └ ∴ ~C {from 6 and 11}
16 ∴ (K ⊃ (U ∨ F)) {from 5; 14 contradicts 15}

15. Invalid

* 1 (A ⊃ B)
* 2 (B ⊃ (F ⊃ M))
* 3 (M ⊃ ~H)
4 H
[∴ ~A
5 asm: A
6 ∴ B {from 1 and 5}
* 7 ∴ (F ⊃ M) {from 2 and 6}
8 ∴ ~M {from 3 and 4}
9 ∴ ~F {from 7 and 8}

An “F” premise would make it valid.

H, A, B, ~M, ~F

0389 20. Valid

* 1 (~A ⊃ (F • C))
* 2 (~G ⊃ ~F)
[∴ (~A ⊃ G)
* 3 ┌ asm: ~(~A ⊃ G)

4 ∴ ~A {from 3}
5 ∴ ~G {from 3}
* 6 ∴ (F • C) {from 1 and 4}
7 ∴ F {from 6}
8 ∴ C {from 6}
9 ∴ ~F {from 2 and 5}
10 ∴ (~A ⊃ G) {from 3; 7 contradicts 9}

25. Valid

* 1 (F ⊃ (E • S))
* 2 (E ⊃ T)
* 3 (T ⊃ T′)
* 4 (T′ ⊃ X)
* 5 ((X • T′) ⊃ Y)
* 6 ((X • Y) ⊃ ~S)
[∴ ~F
7 asm: F
* 8 ∴ (E • S) {from 1 and 7}
9 ∴ E {from 8}
10 ∴ S {from 8}
11 ∴ T {from 2 and 9}
12 ∴ T′ {from 3 and 11}
13 ∴ X {from 4 and 12}
* 14 ∴ ~(X • Y) {from 6 and 10}
15 ∴ ~Y {from 13 and 14}
* 16 ∴ ~(X • T′) {from 5 and 15}
17 ∴ ~X {from 12 and 16}
18 ∴ ~F {from 7; 13 contradicts 17}

7.3a

1. Valid

* 1 (A ⊃ B)
2 (A ∨ (A • C))
[∴ (A • B)
* 3 asm: ~(A • B)
4 asm: ~A {break 1}
5 ∴ (A • C) {from 2 and 4}
6 ∴ A {from 5}
7 ∴ A {from 4; 4 contradicts 6}
8 ∴ B {from 1 and 7}
9 ∴ ~B {from 3 and 7}
10 ∴ (A • B) {from 3; 8 contradicts 9}

3. Valid

* 1 (B ⊃ A)
* 2 ~(A • C)
3 (B ∨ C)
[∴ (A ≡ B)
* 4 asm: ~(A ≡ B)
5 ∴ (A ∨ B) {from 4}

* 6 ∴ ~(A • B) {from 4}
7 asm: ~B {break 1}
8 ∴ C {from 3 and 7}
9 ∴ ~A {from 2 and 8}
10 ∴ A {from 5 and 7}
11 ∴ B {from 7; 9 contradicts 10}
12 ∴ A {from 1 and 11}
13 ∴ ~C {from 2 and 12}
14 ∴ ~A {from 6 and 11}
15 ∴ (A ≡ B) {from 4; 12 contradicts 14}

5. Valid

* 1 ((A ⊃ B) ⊃ C)
* 2 (C ⊃ (D • E))
[∴ (B ⊃ D)
* 3 asm: ~(B ⊃ D)
4 ∴ B {from 3}
5 ∴ ~D {from 3}
6 asm: ~(A ⊃ B) {break 1}
7 ∴ A {from 6}
8 ∴ ~B {from 6}
9 ∴ (A ⊃ B) {from 6; 4 contradicts 8}
10 ∴ C {from 1 and 9}
11 ∴ (D • E) {from 2 and 10}
12 ∴ D {from 11}
13 ∴ (B ⊃ D) {from 3; 5 contradicts 12}

7.3b

1. Valid

* 1 ((F • P) ∨ (A • P))
* 2 (P ⊃ G)
[∴ (P • G)
* 3 asm: ~(P • G)
4 asm: (F • P) {break 1}
5 ∴ F {from 4}
6 ∴ P {from 4}
7 ∴ G {from 2 and 6}
8 ∴ ~G {from 3 and 6}
9 ∴ ~(F • P) {from 4; 7 contradicts 8}
* 10 ∴ (A • P) {from 1 and 9}
11 ∴ A {from 10}
12 ∴ P {from 10}
13 ∴ G {from 2 and 12}
14 ∴ ~G {from 3 and 12}
15 ∴ (P • G) {from 3; 13 contradicts 14}

3. Valid

* 1 (A ≡ (H ∨ B))
2 ~H
[∴ (A ≡ B)

* 3 asm: ∼(A ≡ B)
* 4 ∴ (A ⊃ (H ∨ B)) {from 1}
5 ∴ ((H ∨ B) ⊃ A) {from 1}
6 ∴ (A ∨ B) {from 3}
* 7 ∴ ∼(A • B) {from 3} 0390
8 asm: ∼A {break 4}
9 ∴ ∼(H ∨ B) {from 5 and 8}
10 ∴ ∼B {from 9}
11 ∴ B {from 6 and 8}
12 ∴ A {from 8; 10 contradicts 11}
* 13 ∴ (H ∨ B) {from 4 and 12}
14 ∴ ∼B {from 7 and 12}
15 ∴ B {from 2 and 13}
16 ∴ (A ≡ B) {from 3; 14 contradicts 15}

5. Valid

* 1 (∼C ⊃ ∼M)
* 2 (∼M ⊃ ∼E)
3 (C ⊃ A)
* 4 (A ⊃ E)
[∴ (E ≡ C)
* 5 asm: ∼(E ≡ C)
* 6 ∴ (E ∨ C) {from 5}
7 ∴ ∼(E • C) {from 5}
8 asm: C {break 1}
9 ∴ A {from 3 and 8}
10 ∴ E {from 4 and 9}
11 ∴ M {from 2 and 10}
12 ∴ ∼E {from 7 and 8}
13 ∴ ∼C {from 8; 10 contradicts 12}
14 ∴ ∼M {from 1 and 13}
15 ∴ ∼E {from 2 and 14}
16 ∴ ∼A {from 4 and 15}
17 ∴ E {from 6 and 13}
18 ∴ (E ≡ C) {from 5; 15 contradicts 17}

10. Valid

* 1 ((D • L) ⊃ (F • A))
2 ((D • ∼L) ⊃ (G • A))
[∴ (D ⊃ A)
* 3 asm: ∼(D ⊃ A)
4 ∴ D {from 3}
5 ∴ ∼A {from 3}
6 asm: ∼(D • L) {break 1}
7 ∴ ∼L {from 4 and 6}
8 asm: ∼(D • ∼L) {break 2}
9 ∴ L {from 4 and 8}
10 ∴ (D • ∼L) {from 8; 7 contradicts 9}
11 ∴ (G • A) {from 2 and 10}
12 ∴ G {from 11}
13 ∴ A {from 11}
* 14 ∴ (D • L) {from 6; 5 contradicts 13}

15 | ∴ L {from 14}
* 16 | ∴ (F • A) {from 1 and 14}
17 | ∴ F {from 16}
18 └ ∴ A {from 16}
19 ∴ (D ⊃ A) {from 3; 5 contradicts 18}

7.4a

1. Invalid

1 ~(A • B)
[∴ (~A • ~B)
** 2 asm: ~(~A • ~B)
3 asm: ~A {break 1}
4 ∴ B {from 2 and 3}

~A, B

3. Invalid

1 (A ⊃ B)
2 (C ⊃ (~D • E))
[∴ (D ∨ F)
* 3 asm: ~(D ∨ F)
4 ∴ ~D {from 3}
5 ∴ ~F {from 3}
6 asm: ~A {break 1}
7 asm: ~C {break 2}

~D, ~F, ~A, ~C

5. Invalid

1 (A ⊃ (B • C))
** 2 ((D ⊃ E) ⊃ A)
[∴ (E ∨ C)
* 3 asm: ~(E ∨ C)
4 ∴ ~E {from 3}
5 ∴ ~C {from 3}
6 asm: ~A {break 1}
** 7 ∴ ~(D ⊃ E) {from 2 and 6}
8 ∴ D {from 7}

~E, ~C, ~A, D

7.4b

1. Invalid

1 (M ⊃ ~B)
2 ~M
3 (B ⊃ (P • G))
[∴ G

4 asm: ~G
5 asm: ~B {break 3}

~M, ~G, ~B

3. Invalid

1 (~R ⊃ (O • ~S))
[∴ (R ⊃ (C • S))
* 2 asm: ~(R ⊃ (C • S))
3 ∴ R {from 2}
4 ∴ ~(C • S) {from 2}
5 asm: ~C {break 4}

R, ~C

5. Invalid

1 ((A • L) ⊃ (D • M))
2 (M ⊃ ~C)
3 (S ⊃ L)
[∴ ((A • ~S) ⊃ C)
* 4 asm: ~((A • ~S) ⊃ C)
* 5 ∴ (A • ~S) {from 4}
6 ∴ ~C {from 4} 0391
7 ∴ A {from 5}
8 ∴ ~S {from 5}
** 9 asm: ~(A • L) {break 1}
10 ∴ ~L {from 7 and 9}

~C, A, ~S, ~L

10. Valid

1 H
2 P
3 ~H´
* 4 (M ⊃ ((H • P) ⊃ H´))
5 (~M ⊃ S)
[∴ (~M • S)
6 asm: ~(~M • S)
7 asm: ~M {break 4}
8 ∴ S {from 5 and 7}
9 ∴ ~S {from 6 and 7}
10 ∴ M {from 7; 8 contradicts 9}
* 11 ∴ ((H • P) ⊃ H´) {from 4 and 10}
* 12 ∴ ~(H • P) {from 3 and 11}
13 ∴ ~P {from 1 and 12}
14 ∴ (~M • S) {from 6; 2 contradicts 13}

15. Invalid

1 ((E • F) ⊃ W)
2 ((W • M) ⊃ (B • ~N))
[∴ (N ⊃ ~E)

* 3 asm: ~(N ⊃ ~E)
4 ∴ N {from 3}
5 ∴ E {from 3}
** 6 asm: ~(E • F) {break 1}
7 ∴ ~F {from 5 and 6}
8 asm: ~(W • M) {break 2}
9 asm: ~W {break 8}

N, E, ~F, ~W

Chapter 8 answers

8.1a

1. ~Cx
3. (∃x)~Cx
5. (x)Cx
10. ~(∃x)(Lx • Ex)
15. (∃x)(Ax • (~Bx • Dx))
20. ~(x)(Cx ⊃ Px)
25. (x)(Cx • Lx)

8.2a

1. Valid

1 (x)Fx
[∴ (x)(Gx ∨ Fx)
* 2 ┌ asm: ~(x)(Gx ∨ Fx)
* 3 │ ∴ (∃x)~(Gx ∨ Fx) {from 2}
* 4 │ ∴ ~(Ga ∨ Fa) {from 3}
5 │ ∴ ~Ga {from 4}
6 │ ∴ ~Fa {from 4}
7 └ ∴ Fa {from 1}
8 ∴ (x)(Gx ∨ Fx) {from 2; 6 contradicts 7}

3. Valid

* 1 ~(∃x)(Fx • Gx)
* 2 (∃x)Fx
[∴ (∃x)~Gx
* 3 ┌ asm: ~(∃x)~Gx
4 │ ∴ (x)~(Fx • Gx) {from 1}
5 │ ∴ Fa {from 2}
6 │ ∴ (x)Gx {from 3}
* 7 │ ∴ ~(Fa • Ga) {from 4}
8 │ ∴ ~Ga {from 5 and 7}
9 └ ∴ Ga {from 6}
10 ∴ (∃x)~Gx {from 3; 8 contradicts 9}

5. Valid

1 (x)(Fx ⊃ Gx)
* 2 (∃x)Fx
[∴ (∃x)(Fx • Gx)
* 3 asm: ~(∃x)(Fx • Gx)
4 ∴ Fa {from 2}
5 ∴ (x)~(Fx • Gx) {from 3}
* 6 ∴ (Fa ⊃ Ga) {from 1}
7 ∴ Ga {from 4 and 6}
* 8 ∴ ~(Fa • Ga) {from 5}
9 ∴ ~Ga {from 4 and 8}
10 ∴ (∃x)(Fx • Gx) {from 3; 7 contradicts 9}

10. Valid

1 (x)(Fx ≡ Gx)
* 2 (∃x)~Gx
[∴ (∃x)~Fx
* 3 asm: ~(∃x)~Fx
4 ∴ ~Ga {from 2}
5 ∴ (x)Fx {from 3}
* 6 ∴ (Fa ≡ Ga) {from 1}
* 7 ∴ (Fa ⊃ Ga) {from 6}
8 ∴ (Ga ⊃ Fa) {from 6}
9 ∴ ~Fa {from 4 and 7}
10 ∴ Fa {from 5}
11 ∴ (∃x)~Fx {from 3; 9 contradicts 10}

8.2b

1. Valid

1 (x)(Dx ⊃ Bx)
2 (x)Dx
[∴ (x)Bx
* 3 asm: ~(x)Bx
* 4 ∴ (∃x)~Bx {from 3}
5 ∴ ~Ba {from 4}
* 6 ∴ (Da ⊃ Ba) {from 1}
7 ∴ ~Da {from 5 and 6}
8 ∴ Da {from 2}
9 ∴ (x)Bx {from 3; 7 contradicts 8}

3. Valid

* 1 ~(∃x)(Fx • Ox) 0392
2 (x)(Cx ⊃ Ox)
[∴ ~(∃x)(Fx • Cx)
* 3 asm: (∃x)(Fx • Cx)
4 ∴ (x)~(Fx • Ox) {from 1}
* 5 ∴ (Fa • Ca) {from 3}
6 ∴ Fa {from 5}
7 ∴ Ca {from 5}
* 8 ∴ (Ca ⊃ Oa) {from 2}

```
   9 │ ∴ Oa    {from 7 and 8}
* 10 │ ∴ ~(Fa • Oa)    {from 4}
  11 └ ∴ ~Oa    {from 6 and 10}
  12 ∴ ~(∃x)(Fx • Cx)    {from 3; 9 contradicts 11}
```

5. Valid

```
    1   (x)(Kx ⊃ Ex)
*   2   ~(∃x)(Ex • Kx)
    [ ∴ ~(∃x)Kx
*   3 ┌ asm: (∃x)Kx
    4 │ ∴ (x)~(Ex • Kx)    {from 2}
    5 │ ∴ Ka    {from 3}
*   6 │ ∴ (Ka ⊃ Ea)    {from 1}
    7 │ ∴ Ea    {from 5 and 6}
*   8 │ ∴ ~(Ea • Ka)    {from 4}
    9 └ ∴ ~Ea    {from 5 and 8}
   10 ∴ ~(∃x)Kx    {from 3; 7 contradicts 9}
```

10. Valid

```
*   1   ~(∃x)(Bx • Cx)
*   2   (∃x)(Lx • Cx)
    3   (x)(Cx ⊃ Rx)
    [ ∴ (∃x)(Rx • (Cx • ~Bx))
*   4 ┌ asm: ~(∃x)(Rx • (Cx • ~Bx))
    5 │ ∴ (x)~(Bx • Cx)    {from 1}
*   6 │ ∴ (La • Ca)    {from 2}
    7 │ ∴ (x)~(Rx • (Cx • ~Bx))    {from 4}
    8 │ ∴ La    {from 6}
    9 │ ∴ Ca    {from 6}
*  10 │ ∴ (Ca ⊃ Ra)    {from 3}
   11 │ ∴ Ra    {from 9 and 10}
*  12 │ ∴ ~(Ba • Ca)    {from 5}
   13 │ ∴ ~Ba    {from 9 and 12}
*  14 │ ∴ ~(Ra • (Ca • ~Ba))    {from 7}
*  15 │ ∴ ~(Ca • ~Ba)    {from 11 and 14}
   16 └ ∴ Ba    {from 9 and 15}
   17 ∴ (∃x)(Rx • (Cx • ~Bx))    {from 4; 13 contradicts 16}
```

8.3a

1. Invalid

```
* 1   (∃x)Fx
  [ ∴ (x)Fx
* 2   asm: ~(x)Fx
  3   ∴ Fa    {from 1}
* 4   ∴ (∃x)~Fx    {from 2}
  5   ∴ ~Fb    {from 4}
```

a, b

Fa, ~Fb

3. Invalid

* 1 (∃x)(Fx ∨ Gx)
* 2 ~(x)Fx
[∴ (∃x)Gx
* 3 asm: ~(∃x)Gx
* 4 ∴ (Fa ∨ Ga) {from 1}
* 5 ∴ (∃x)~Fx {from 2}
6 ∴ (x)~Gx {from 3}
7 ∴ ~Fb {from 5}
8 ∴ ~Ga {from 6}
9 ∴ Fa {from 4 and 8}
10 ∴ ~Gb {from 6}

a, b

Fa, ~Ga, ~Fb, ~Gb

5. Invalid

* 1 ~(∃x)(Fx • Gx)
2 (x)~Fx
[∴ (x)Gx
* 3 asm: ~(x)Gx
4 ∴ (x)~(Fx • Gx) {from 1}
* 5 ∴ (∃x)~Gx {from 3}
6 ∴ ~Ga {from 5}
7 ∴ ~Fa {from 2}
8 ∴ ~(Fa • Ga) {from 4}

a

~Ga, ~Fa

10. Invalid

* 1 (∃x)~Fx
* 2 (∃x)~Gx
[∴ (∃x)(Fx ≡ Gx)
* 3 asm: ~(∃x)(Fx ≡ Gx)
4 ∴ ~Fa {from 1}
5 ∴ ~Gb {from 2}
6 ∴ (x)~(Fx ≡ Gx) {from 3}
* 7 ∴ ~(Fa ≡ Ga) {from 6}
* 8 ∴ (Fa ∨ Ga) {from 7}
9 ∴ ~(Fa • Ga) {from 7}
10 ∴ Ga {from 4 and 8}
* 11 ∴ ~(Fb ≡ Gb) {from 6}
* 12 ∴ (Fb ∨ Gb) {from 11}
13 ∴ ~(Fb • Gb) {from 11}
14 ∴ Fb {from 5 and 12}

a, b

Ga, ~Fa, Fb, ~Gb

8.3b

1. Invalid

* 1 (∃x)(Bx • Gx)
[∴ (x)(Bx ⊃ Gx)
* 2 asm: ~(x)(Bx ⊃ Gx)
* 3 ∴ (∃x)~(Bx ⊃ Gx) {from 2}
* 4 ∴ (Ba • Ga) {from 1}
5 ∴ Ba {from 4}
6 ∴ Ga {from 4}
* 7 ∴ ~(Bb ⊃ Gb) {from 3}
8 ∴ Bb {from 7}
9 ∴ ~Gb {from 7}

a, b

Ba, Ga, Bb, ~Gb

0393 3. Invalid

* 1 (∃x)Sx
* 2 ~(x)Cx
[∴ (∃x)(Sx • ~Cx)
* 3 asm: ~(∃x)(Sx • ~Cx)
4 ∴ Sa {from 1}
* 5 ∴ (∃x)~Cx {from 2}
6 ∴ (x)~(Sx • ~Cx) {from 3}
7 ∴ ~Cb {from 5}
* 8 ∴ ~(Sa • ~Ca) {from 6}
9 ∴ Ca {from 4 and 8}
* 10 ∴ ~(Sb • ~Cb) {from 6}
11 ∴ ~Sb {from 7 and 10}

a, b

Sa, Ca, ~Sb, ~Cb

5. Valid

1 (x)((Vx • Cx) ⊃ Px)
2 (x)(Dx ⊃ (Cx • ~Px))
[∴ ~(∃x)(Dx • Vx)
* 3 ┌ asm: (∃x)(Dx • Vx)
* 4 │ ∴ (Da • Va) {from 3}
5 │ ∴ Da {from 4}
6 │ ∴ Va {from 4}
* 7 │ ∴ ((Va • Ca) ⊃ Pa) {from 1}
* 8 │ ∴ (Da ⊃ (Ca • ~Pa)) {from 2}
* 9 │ ∴ (Ca • ~Pa) {from 5 and 8}
10 │ ∴ Ca {from 9}
11 │ ∴ ~Pa {from 9}
* 12 │ ∴ ~(Va • Ca) {from 7 and 11}
13 └ ∴ ~Ca {from 6 and 12}
14 ∴ ~(∃x)(Dx • Vx) {from 3; 10 contradicts 13}

10. Valid

1 (x)(Sx ⊃ Vx)
[∴ (x)(~Vx ⊃ ~Sx)
* 2 asm: ~(x)(~Vx ⊃ ~Sx)
* 3 ∴ (∃x)~(~Vx ⊃ ~Sx) {from 2}
* 4 ∴ ~(~Va ⊃ ~Sa) {from 3}
5 ∴ ~Va {from 4}
6 ∴ Sa {from 4}
* 7 ∴ (Sa ⊃ Va) {from 1}
8 ∴ ~Sa {from 5 and 7}
9 ∴ (x)(~Vx ⊃ ~Sx) {from 2; 6 contradicts 8}

15. Invalid

* 1 ~(∃x)(Px • Bx)
* 2 (∃x)(Cx • ~Bx)
[∴ (∃x)(Cx • Px)
* 3 asm: ~(∃x)(Cx • Px)
4 ∴ (x)~(Px • Bx) {from 1}
* 5 ∴ (Ca • ~Ba) {from 2}
6 ∴ (x)~(Cx • Px) {from 3}
7 ∴ Ca {from 5}
8 ∴ ~Ba {from 5}
9 ∴ ~(Pa • Ba) {from 4}
* 10 ∴ ~(Ca • Pa) {from 6}
11 ∴ ~Pa {from 7 and 10}

a
Ca, ~Ba, ~Pa

8.4a

1. (Cg ∨ Eg)
3. ((x)Lx ⊃ (x)Ex)
5. ((∃x)Ex ⊃ R)
10. ((x)Ex ⊃ (x)(Lx ⊃ Ex))
15. ~(∃x)Ex *or, equivalently,* (x)~Ex
20. ~(∃x)(Lx • Ex) *or, equiv,* (x)~(Lx • Ex)

8.5a

1. Valid

1 (x)(Fx ∨ Gx)
2 ~Fa
[∴ (∃x)Gx
* 3 asm: ~(∃x)Gx
4 ∴ (x)~Gx {from 3}
* 5 ∴ (Fa ∨ Ga) {from 1}
6 ∴ Ga {from 2 and 5}
7 ∴ ~Ga {from 4}
8 ∴ (∃x)Gx {from 3; 6 contradicts 7}

3. Invalid

* 1 ((x)Ex ⊃ R)
[∴ (x)(Ex ⊃ R)
* 2 asm: ~(x)(Ex ⊃ R)
* 3 ∴ (∃x)~(Ex ⊃ R) {from 2}
* 4 ∴ ~(Ea ⊃ R) {from 3}
5 ∴ Ea {from 4}
6 ∴ ~R {from 4}
* 7 ∴ ~(x)Ex {from 1 and 6}
* 8 ∴ (∃x)~Ex {from 7}
9 ∴ ~Eb {from 8}

a, b

Ea, ~Eb, ~R

5. Invalid

* 1 ((∃x)Fx ⊃ (∃x)Gx)
[∴ (x)(Fx ⊃ Gx)
* 2 asm: ~(x)(Fx ⊃ Gx)
* 3 ∴ (∃x)~(Fx ⊃ Gx) {from 2}
* 4 ∴ ~(Fa ⊃ Ga) {from 3}
5 ∴ Fa {from 4}
6 ∴ ~Ga {from 4}
7 asm: ~(∃x)Fx {break 1}
8 ∴ (x)~Fx {from 7}
9 ∴ ~Fa {from 8}
10 ∴ (∃x)Fx {from 7; 5 contradicts 9}
* 11 ∴ (∃x)Gx {from 1 and 10}
12 ∴ Gb {from 11}

a, b

Fa, ~Ga, Gb

10. Invalid

* 1 ~(∃x)(Fx • Gx)
2 ~Fd
[∴ Gd
3 asm: ~Gd
4 ∴ (x)~(Fx • Gx) {from 1}
5 ∴ ~(Fd • Gd) {from 4}

d

~Fd, ~Gd

0394 15. Valid

* 1 ((∃x)Ex ⊃ R)
[∴ (x)(Ex ⊃ R)
* 2 asm: ~(x)(Ex ⊃ R)
* 3 ∴ (∃x)~(Ex ⊃ R) {from 2}

* 4 | ∴ ~(Ea ⊃ R) {from 3}
5 | ∴ Ea {from 4}
6 | ∴ ~R {from 4}
* 7 | ∴ ~(∃x)Ex {from 1 and 6}
8 | ∴ (x)~Ex {from 7}
9 └ ∴ ~Ea {from 8}
10 ∴ (x)(Ex ⊃ R) {from 2; 5 contradicts 9}

8.5b

1. Valid

1 (x)Cx
* 2 (Cw ⊃ G)
[∴ G
3 ┌ asm: ~G
4 | ∴ ~Cw {from 2 and 3}
5 └ ∴ Cw {from 1}
6 ∴ G {from 3; 4 contradicts 5}

3. Valid

1 (x)(Bx ⊃ Cx)
2 Bw
* 3 (Cw ⊃ G)
[∴ G
4 ┌ asm: ~G
5 | ∴ ~Cw {from 3 and 4}
* 6 | ∴ (Bw ⊃ Cw) {from 1}
7 └ ∴ Cw {from 2 and 6}
8 ∴ G {from 4; 5 contradicts 7}

5. Valid

1 (x)(Ex ⊃ (Ix ∨ Fx))
[∴ (~(∃x)Ix ⊃ (x)(~Fx ⊃ ~Ex))
* 2 ┌ asm: ~(~(∃x)Ix ⊃ (x)(~Fx ⊃ ~Ex))
* 3 | ∴ ~(∃x)Ix {from 2}
* 4 | ∴ ~(x)(~Fx ⊃ ~Ex) {from 2}
5 | ∴ (x)~Ix {from 3}
* 6 | ∴ (∃x)~(~Fx ⊃ ~Ex) {from 4}
* 7 | ∴ ~(~Fa ⊃ ~Ea) {from 6}
8 | ∴ ~Fa {from 7}
9 | ∴ Ea {from 7}
* 10 | ∴ (Ea ⊃ (Ia ∨ Fa)) {from 1}
* 11 | ∴ (Ia ∨ Fa) {from 9 and 10}
12 | ∴ Ia {from 8 and 11}
13 └ ∴ ~Ia {from 5}
14 ∴ (~(∃x)Ix ⊃ (x)(~Fx ⊃ ~Ex)) {from 2; 12 contradicts 13}

10. Valid

1 (x)(Ax ⊃ Px)
* 2 (Ae • Ad)

```
    [ ∴ (Pe • Pd)
*   3  ┌ asm: ~(Pe • Pd)
    4  │ ∴ Ae    {from 2}
    5  │ ∴ Ad    {from 2}
*   6  │ ∴ (Ad ⊃ Pd)    {from 1}
    7  │ ∴ Pd    {from 5 and 6}
    8  │ ∴ ~Pe    {from 3 and 7}
*   9  │ ∴ (Ae ⊃ Pe)    {from 1}
   10  └ ∴ Pe    {from 4 and 9}
   11 ∴ (Pe • Pd)    {from 3; 8 contradicts 10}
```

15. Valid

```
*   1  (T ⊃ (H • (x)(Mx ⊃ Ex)))
    2  (x)(Ex ⊃ Ix)
    3  Mt
    4  ~It
    [ ∴ ~T
    5  ┌ asm: T
*   6  │ ∴ (H • (x)(Mx ⊃ Ex))    {from 1 and 5}
    7  │ ∴ H    {from 6}
    8  │ ∴ (x)(Mx ⊃ Ex)    {from 6}
*   9  │ ∴ (Et ⊃ It)    {from 2}
   10  │ ∴ ~Et    {from 4 and 9}
*  11  │ ∴ (Mt ⊃ Et)    {from 8}
   12  └ ∴ Et    {from 3 and 11}
   13 ∴ ~T    {from 5; 10 contradicts 12}
```

20. Valid

```
*   1  ~(∃x)(~Cx • Ix)
    2  (x)(Ex ⊃ Ix)
    [ ∴ (x)(Ex ⊃ Cx)
*   3  ┌ asm: ~(x)(Ex ⊃ Cx)
    4  │ ∴ (x)~(~Cx • Ix)    {from 1}
*   5  │ ∴ (∃x)~(Ex ⊃ Cx)    {from 3}
*   6  │ ∴ ~(Ea ⊃ Ca)    {from 5}
    7  │ ∴ Ea    {from 6}
    8  │ ∴ ~Ca    {from 6}
*   9  │ ∴ (Ea ⊃ Ia)    {from 2}
   10  │ ∴ Ia    {from 7 and 9}
*  11  │ ∴ ~(~Ca • Ia)    {from 4}
   12  └ ∴ ~Ia    {from 8 and 11}
   13 ∴ (x)(Ex ⊃ Cx)    {from 3; 10 contradicts 12}
```

Chapter 9 answers

9.1a

1. La
3. ~a=p

5. (∃x)(∃y)(~x=y • (Lx • Ly))
10. (∃x)(Lx • ~(∃y)(~y=x • Ly))
15. (Ra • ~a=f)

9.2a

1. Invalid

  1   Fa
  [∴ ~(∃x)(Fx • ~x=a) 0395
* 2   asm: (∃x)(Fx • ~x=a)
* 3   ∴ (Fb • ~b=a)   {from 2}
  4   ∴ Fb   {from 3}
  5   ∴ ~b=a   {from 3}

a, b

Fa, ~Fb, ~b=a

3. Valid

1   a=b
2   b=c
[∴ a=c
3 ┌ asm: ~a=c
4 └ ∴ ~b=c   {from 1 and 3}
5 ∴ a=c   {from 3; 2 contradicts 4}

5. Invalid

1   ~a=b
2   ~c=b
[∴ a=c
3   asm: ~a=c

a, b, c

~a=b, ~a=c, ~c=b

10. Invalid

  [∴ (∃x)(∃y)~y=x
* 1   asm: ~(∃x)(∃y)~y=x
  2   ∴ (x)~(∃y)~y=x   {from 1}
* 3   ∴ ~(∃y)~y=a   {from 2}
  4   ∴ (y)y=a   {from 3}
  5   ∴ a=a   {from 4}

a

a=a

9.2b

1. Valid

```
  1   k=n
  2   Bn
  3   Tk
  [ ∴ (∃x)(Tx • Bx)
* 4 ┌ asm: ~(∃x)(Tx • Bx)
  5 │ ∴ Bk    {from 1 and 2}
  6 │ ∴ (x)~(Tx • Bx)    {from 4}
* 7 │ ∴ ~(Tk • Bk)    {from 6}
  8 └ ∴ ~Bk    {from 3 and 7}
  9 ∴ (∃x)(Tx • Bx)    {from 4; 5 contradicts 8}
```

3. Valid

```
1   Oc
2   ~Op
[ ∴ ~p=c
3 ┌ asm: p=c
4 └ ∴ ~Oc    {from 2 and 3}
5 ∴ ~p=c    {from 3; 1 contradicts 4}
```

5. Invalid

```
1   ~Bm
2   ~Bu
[ ∴ u=m
3   asm: ~u=m
```

m, u

~Bm, ~Bu, ~u=m

10. Valid

```
* 1   (Ku ∨ Kt)
  2   ~Ku
  [ ∴ ~u=t
  3 ┌ asm: u=t
  4 │ ∴ Kt    {from 1 and 2}
  5 └ ∴ ~Kt    {from 2 and 3}
  6 ∴ ~u=t    {from 3; 4 contradicts 5}
```

15. Valid

```
*  1   (∃x)(Lx • ~(∃y)(~y=x • Ly))
   2   Lp
   3   Fp
   [ ∴ (x)(Lx ⊃ Fx)
*  4 ┌ asm: ~(x)(Lx ⊃ Fx)
*  5 │ ∴ (∃x)~(Lx ⊃ Fx)    {from 4}
*  6 │ ∴ ~(La ⊃ Fa)    {from 5}
   7 │ ∴ La    {from 6}
```

8 | ∴ ~Fa {from 6}
* 9 | ∴ (Lb • ~(∃y)(~y=b • Ly)) {from 1}
10 | ∴ Lb {from 9}
* 11 | ∴ ~(∃y)(~y=b • Ly) {from 9}
12 | ∴ (y)~(~y=b • Ly) {from 11}
* 13 | ∴ ~(~a=b • La) {from 12}
14 | ∴ a=b {from 7 and 13}
* 15 | ∴ ~(~p=b • Lp) {from 12}
16 | ∴ p=b {from 2 and 15}
17 | ∴ ~Fb {from 8 and 14}
18 └ ∴ ~Fp {from 16 and 17}
19 ∴ (x)(Lx ⊃ Fx) {from 4; 3 contradicts 18}

9.3a

1. (Lto • Lot)
3. (x)(Rx ⊃ Ltx)
5. ((x)Lxo • ~(x)Lox)
10. (x)(Lxx ⊃ Lox)
15. (x)(Cgx ⊃ Lgx)
20. (x)(Cgx ⊃ Ggx)

9.4a

1. (x)(Rx ⊃ (y)Lyx) *or, equivalently,* (x)(y)(Ry ⊃ Lxy)
3. (∃x)(Rx • (∃y)Lyx) *or, equivalently,* (∃x)(∃y)(Ry • Lxy)
5. (x)(Rx ⊃ (∃y)(Iy • Lxy))
10. ~(∃x)(Ix • (y)Lxy)
15. ((x)Ltx ⊃ (∃x)(Ix • (y)Lxy))
20. (x)(∃y)Cyx
25. (x)(y)(Cxy ⊃ Lxy)

9.5a

1. Invalid

1 (x)Lxa
[∴ (x)Lax
* 2 asm: ~(x)Lax 0396
* 3 ∴ (∃x)~Lax {from 2}
4 ∴ ~Lab {from 3}
5 ∴ Laa {from 1}
6 ∴ Lba {from 1}

a, b

Lab, Laa, ~Lab

3. Invalid

1 (x)(y)(Lxy ⊃ x=y)
[∴ (x)Lxx

* 2 asm: ~(x)Lxx
* 3 ∴ (∃x)~Lxx {from 2}
4 ∴ ~Laa {from 3}
5 ∴ (y)(Lay ⊃ a=y) {from 1}
6 ∴ (Laa ⊃ a=a) {from 5}

a

~Laa

5. Valid

1 (x)(y)Lxy
[∴ (x)(y)((Fx • Gy) ⊃ Lxy)
* 2 ┌ asm: ~(x)(y)((Fx • Gy) ⊃ Lxy)
* 3 │ ∴ (∃x)~(y)((Fx • Gy) ⊃ Lxy) {from 2}
* 4 │ ∴ ~(y)((Fa • Gy) ⊃ Lay) {from 3}
* 5 │ ∴ (∃y)~((Fa • Gy) ⊃ Lay) {from 4}
* 6 │ ∴ ~((Fa • Gb) ⊃ Lab) {from 5}
* 7 │ ∴ (Fa • Gb) {from 6}
8 │ ∴ ~Lab {from 6}
9 │ ∴ Fa {from 7}
10 │ ∴ Gb {from 7}
11 │ ∴ (y)Lay {from 1}
12 │ ∴ (y)Lby {from 1}
13 │ ∴ Laa {from 11}
14 └ ∴ Lab {from 11}
15 ∴ (x)(y)((Fx • Gy) ⊃ Lxy) {from 2; 8 contradicts 14}

10. Valid

1 Lab
2 Lbc
[∴ (∃x)(Lax • Lxc)
* 3 ┌ asm: ~(∃x)(Lax • Lxc)
4 │ ∴ (x)~(Lax • Lxc) {from 3}
5 │ ∴ ~(Laa • Lac) {from 4}
* 6 │ ∴ ~(Lab • Lbc) {from 4}
7 └ ∴ ~Lbc {from 1 and 6}
8 ∴ (∃x)(Lax • Lxc) {from 3; 2 contradicts 7}

15. Valid

1 (x)(y)(Lxy ⊃ (Fx • ~Fy))
[∴ (x)(y)(Lxy ⊃ ~Lyx)
* 2 ┌ asm: ~(x)(y)(Lxy ⊃ ~Lyx)
* 3 │ ∴ (∃x)~(y)(Lxy ⊃ ~Lyx) {from 2}
* 4 │ ∴ ~(y)(Lay ⊃ ~Lya) {from 3}
* 5 │ ∴ (∃y)~(Lay ⊃ ~Lya) {from 4}
* 6 │ ∴ ~(Lab ⊃ ~Lba) {from 5}
7 │ ∴ Lab {from 6}
8 │ ∴ Lba {from 6}
9 │ ∴ (y)(Lay ⊃ (Fa • ~Fy)) {from 1}
10 │ ∴ (y)(Lby ⊃ (Fb • ~Fy)) {from 1}
* 11 │ ∴ (Lab ⊃ (Fa • ~Fb)) {from 9}

* 12 ∴ (Fa • ~Fb)   {from 7 and 11}
13 ∴ Fa   {from 12}
14 ∴ ~Fb   {from 12}
* 15 ∴ (Lba ⊃ (Fb • ~Fa))   {from 10}
16 ∴ (Fb • ~Fa)   {from 8 and 15}
17 ∴ Fb   {from 16}
18 ∴ (x)(y)(Lxy ⊃ ~Lyx)   {from 2; 14 contradicts 17}

9.5b

1. Valid

1 (x)Ljx
[∴ (∃x)Lxu
* 2 asm: ~(∃x)Lxu
3 ∴ (x)~Lxu   {from 2}
4 ∴ Ljj   {from 1}
5 ∴ Lju   {from 1}
6 ∴ ~Lju   {from 3}
7 ∴ (∃x)Lxu   {from 2; 5 contradicts 6}

3. Invalid

1 Oab
[∴ ~Oba
2 asm: Oba

a, b

Oab, Oba

To make it valid, we need the premise that "older than" is asymmetrical: "(x)(y)(Oxy ⊃ ~Oyx)" –"In every case, if x is older than y, then y isn't older than x."

5. Invalid

1 (x)(∃y)Dxy
[∴ (∃y)(x)Dxy
* 2 asm: ~(∃y)(x) Dxy
3 ∴ (y)~(x)Dxy   {from 2}
* 4 ∴ (∃y)Day   {from 1}
5 ∴ Dab   {from 4}
6 ∴ ~(x)Dxb   {from 3}
7 ∴ (∃x)~Dxb   {from 6}

Endless loop: we add further wffs to make the premise true and conclusion false. "~Dab, ~Dba, Daa, Dbb" also refutes the argument.

a, b

Dab, Dba, ~Daa, ~Dbb

10. Valid

* 1 (∃x)(y)Lyx
[∴ (∃x)Lxx

* 2 ┌ asm: ~(∃x)Lxx
3 │ ∴ (y)Lya {from 1}
4 │ ∴ (x)~Lxx {from 2}
5 │ ∴ Laa {from 3}
6 └ ∴ ~Laa {from 4}
7 ∴ (∃x)Lxx {from 2; 5 contradicts 6}

15. Valid

1 (x)((∃y)Lxy ⊃ (y)Lyx)
2 Lrj 0397
[∴ Liu
3 ┌ asm: ~Liu
* 4 │ ∴ ((∃y)Lry ⊃ (y)Lyr) {from 1}
5 │ ┌ asm: ~(∃y)Lry {break 4}
6 │ │ ∴ (y)~Lry {from 5}
7 │ └ ∴ ~Lrj {from 6}
8 │ ∴ (∃y)Lry {from 5; 2 contradicts 7}
9 │ ∴ (y)Lyr {from 4 and 8}
10 │ ∴ Lur {from 9}
* 11 │ ∴ ((∃y)Luy ⊃ (y)Lyu) {from 1}
12 │ ┌ asm: ~(∃y)Luy {break 11}
13 │ │ ∴ (y)~Luy {from 12}
14 │ └ ∴ ~Lur {from 13}
15 │ ∴ (∃y)Luy {from 12; 10 contradicts 14}
16 │ ∴ (y)Lyu {from 11 and 15}
17 └ ∴ Liu {from 16}
18 ∴ Liu {from 3; 3 contradicts 17}

20. Valid

1 (x)(Gx ⊃ (y)((Ey • Cxy) ⊃ Pxy))
2 (x)(Ox ⊃ (y)(Ey ⊃ Cxy))
3 ((∃x)(y)(Ey ⊃ Pxy) ⊃ ~(∃x)Ex)
* 4 (∃x)Ex
[∴ (~Og ∨ ~Gg)
* 5 ┌ asm: ~(~Og ∨ ~Gg)
6 │ ∴ Og {from 5}
7 │ ∴ Gg {from 5}
* 8 │ ∴ (Gg ⊃ (y)((Ey • Cgy) ⊃ Pgy)) {fm 1}
9 │ ∴ (y)((Ey • Cgy) ⊃ Pgy) {fm 7 and 8}
* 10 │ ∴ (Og ⊃ (y)(Ey ⊃ Cgy)) {from 2}
11 │ ∴ (y)(Ey ⊃ Cgy) {from 7 and 8}
12 │ ∴ Ea {from 4}
* 13 │ ∴ ~(∃x)(y)(Ey ⊃ Pxy) {from 3 and 4}
14 │ ∴ (x)~(y)(Ey ⊃ Pxy) {from 13}
* 15 │ ∴ ~(y)(Ey ⊃ Pgy) {from 14}
* 16 │ ∴ (∃y)~(Ey ⊃ Pgy) {from 15}
* 17 │ ∴ ~(Eb ⊃ Pgb) {from 16}
18 │ ∴ Eb {from 17}
19 │ ∴ ~Pgb {from 17}
* 20 │ ∴ ((Eb • Cgb) ⊃ Pgb) {from 9}
* 21 │ ∴ (Eb ⊃ Cgb) {from 11}
22 │ ∴ Cgb {from 18 and 21}

* 23 | ∴ ~(Eb • Cgb) {from 19 and 20}
24 └ ∴ ~Eb {from 22 and 23}
25 ∴ (~Og ∨ ~Gg) {from 5; 18 contradicts 24}

25. Valid

1 (x)(Dx ⊃ Ax)
[∴ (x)((∃y)(Dy • Hxy) ⊃ (∃y)(Ay • Hxy))
* 2 ┌ asm: ~(x)((∃y)(Dy • Hxy) ⊃ (∃y)(Ay • Hxy))
* 3 | ∴ (∃x)~((∃y)(Dy • Hxy) ⊃ (∃y)(Ay • Hxy)) {from 2}
* 4 | ∴ ~((∃y)(Dy • Hay) ⊃ (∃y)(Ay • Hay)) {from 3}
* 5 | ∴ (∃y)(Dy • Hay) {from 4}
* 6 | ∴ ~(∃y)(Ay • Hay) {from 4}
* 7 | ∴ (Db • Hab) {from 5}
8 | ∴ Db {from 7}
9 | ∴ Hab {from 7}
10 | ∴ (y)~(Ay • Hay) {from 6}
* 11 | ∴ ~(Ab • Hab) {from 10}
12 | ∴ ~Ab {from 9 and 11}
* 13 | ∴ (Db ⊃ Ab) {from 1}
14 └ ∴ Ab {from 8 and 13}
15 ∴ (x)((∃y)(Dy • Hxy) ⊃ (∃y)(Ay • Hxy)) {from 2; 12 contradicts 14}

Chapter 10 answers

10.1a

1. ☐G
3. ~☐M
5. ☐(R ⊃ P)
10. Ambiguous: (R ⊃ ☐R) or ☐(R ⊃ R)
15. (A ⊃ ☐B)
20. ☐(H ∨ T)
25. (R ⊃ ☐E)
30. ☐(G ⊃ ☐G)

10.2a

1. Valid

* 1 ◇(A • B)
[∴ ◇A
* 2 ┌ asm: ~◇A
* 3 | W ∴ (A • B) {from 1}
4 | ∴ ☐~A {from 2}
5 | W ∴ A {from 3}
6 | W ∴ B {from 3}
7 └ W ∴ ~A {from 4}
8 ∴ ◇A {from 2; 5 contradicts 7}

3. Valid

```
*  1   ~◇(A • ~B)
       [ ∴ □(A ⊃ B)
*  2 ┌ asm: ~□(A ⊃ B)
   3 │ ∴ □~(A • ~B)    {from 1}
*  4 │ ∴ ◇~(A ⊃ B)    {from 2}
*  5 │ W ∴ ~(A ⊃ B)    {from 4}
   6 │ W ∴ A    {from 5}
   7 │ W ∴ ~B    {from 5}
*  8 │ W ∴ ~(A • ~B)    {from 3}
   9 └ W ∴ B    {from 6 and 8}
  10 ∴ □(A ⊃ B)    {from 2; 7 contradicts 9}
```

5. Valid

```
*  1   (◇A ∨ ◇B)
       [ ∴ ◇(A ∨ B)
*  2 ┌ asm: ~◇(A ∨ B)
   3 │ ∴ □~(A ∨ B)    {from 2}
   4 │ ┌ asm: ◇A    {break 1}
   5 │ │ W ∴ A    {from 4}
   6 │ │ W ∴ ~(A ∨ B)    {from 3} 0398
   7 │ └ W ∴ ~A    {from 6}
*  8 │ ∴ ~◇A    {from 4; 5 contradicts 7}
   9 │ ∴ □~A    {from 8}
* 10 │ ∴ ◇B    {from 1 and 8}
  11 │ WW ∴ B    {from 10}
* 12 │ WW ∴ ~(A ∨ B)    {from 3}
  13 │ WW ∴ ~A    {from 12}
  14 └ WW ∴ ~B    {from 12}
  15 ∴ ◇(A ∨ B)    {from 2; 11 contradicts 14}
```

10. Valid

```
   1   □(A ⊃ B)
       [ ∴ (□A ⊃ □B)
*  2 ┌ asm: ~(□A ⊃ □B)
   3 │ ∴ □A    {from 2}
*  4 │ ∴ ~□B    {from 2}
*  5 │ ∴ ◇~B    {from 4}
   6 │ W ∴ ~B    {from 5}
*  7 │ W ∴ (A ⊃ B)    {from 1}
   8 │ W ∴ ~A    {from 6 and 7}
   9 └ W ∴ A    {from 3}
  10 ∴ (□A ⊃ □B)    {from 2; 8 contradicts 9}
```

10.2b

1. Valid

```
   1   □(T ⊃ L)
```

* 2 ◇(T • ~I)
[∴ ◇(L • ~I)
* 3 ┌ asm: ~◇(L • ~I)
* 4 │ W ∴ (T • ~I) {from 2}
5 │ ∴ ☐~(L • ~I) {from 3}
6 │ W ∴ T {from 4}
7 │ W ∴ ~I {from 4}
* 8 │ W ∴ (T ⊃ L) {from 1}
9 │ W ∴ L {from 6 and 8}
* 10 │ W ∴ ~(L • ~I) {from 5}
11 └ W ∴ ~L {from 7 and 10}
12 ∴ ◇(L • ~I) {from 3; 9 contradicts 11}

3. Valid

* 1 (C ⊃ ☐(T ⊃ M))
* 2 ~☐M
3 ☐(~T ⊃ ~~T)
[∴ ~C
4 ┌ asm: C
5 │ ∴ ☐(T ⊃ M) {from 1 and 4}
* 6 │ ∴ ◇~M {from 2}
7 │ W ∴ ~M {from 6}
* 8 │ W ∴ (~T ⊃ ~~T) {from 3}
* 9 │ W ∴ (T ⊃ M) {from 5}
10 │ W ∴ ~T {from 7 and 9}
11 └ W ∴ ~~T {from 8 and 10}
12 ∴ ~C {from 4; 10 contradicts 11}

5. Valid

* 1 ◇(G • T)
2 ☐(T ⊃ E)
[∴ ◇(G • E)
* 3 ┌ asm: ~◇(G • E)
* 4 │ W ∴ (G • T) {from 1}
5 │ ∴ ☐~(G • E) {from 3}
6 │ W ∴ G {from 4}
7 │ W ∴ T {from 4}
* 8 │ W ∴ (T ⊃ E) {from 2}
9 │ W ∴ E {from 7 and 8}
* 10 │ W ∴ ~(G • E) {from 5}
11 └ W ∴ ~E {from 6 and 10}
12 ∴ ◇(G • E) {from 3; 9 contradicts 11}

10. Valid

1 ☐((R • B) ⊃ ~N)
* 2 ◇(R • B)
[∴ ~☐(R ⊃ N)
3 ┌ asm: ☐(R ⊃ N)
* 4 │ W ∴ (R • B) {from 2}

5 │ W ∴ R {from 4}
6 │ W ∴ B {from 4}
* 7 │ W ∴ ((R • B) ⊃ ~N) {from 1}
8 │ W ∴ ~N {from 4 and 7}
* 9 │ W ∴ (R ⊃ N) {from 3}
10 └ W ∴ N {from 5 and 9}
11 ∴ ~☐(R ⊃ N) {from 3; 8 contradicts 10}

15. Valid

1 ☐(~T ⊃ T)
[∴ ☐T
* 2 ┌ asm: ~☐T
* 3 │ ∴ ◇~T {from 2}
4 │ W ∴ ~T {from 3}
* 5 │ W ∴ (~T ⊃ T) {from 1}
6 └ W ∴ T {from 4 and 5}
7 ∴ ☐T {from 2; 4 contradicts 6}

10.3a

1. Invalid

* 1 ◇A
[∴ ☐A
* 2 asm: ~☐A
3 W ∴ A {from 1}
* 4 ∴ ◇~A {from 2}
5 WW ∴ ~A {from 4}

W	A
WW	~A

3. Invalid

* 1 ◇A
* 2 ◇B
[∴ ◇(A • B)
* 3 asm: ~◇(A • B)
4 W ∴ A {from 1}
5 WW ∴ B {from 2}
6 ∴ ☐~(A • B) {from 3}
* 7 W ∴ ~(A • B) {from 6}
8 W ∴ ~B {from 4 and 7} 0399
* 9 WW ∴ ~(A • B) {from 6}
10 WW ∴ ~A {from 5 and 9}

W	A, ~B
WW	B, ~A

5. Invalid

```
    1  (☐A ⊃ ☐B)
     [ ∴ ☐(A ⊃ B)
*   2  asm: ~☐(A ⊃ B)
*   3  ∴ ◇~(A ⊃ B)    {from 2}
*   4  W ∴ ~(A ⊃ B)    {from 3}
    5  W ∴ A    {from 4}
    6  W ∴ ~B    {from 4}
**  7    asm: ~☐A    {break 1}
**  8    ∴ ◇~A    {from 7}
    9    WW ∴ ~A    {from 8}
```

W	A, ~B
WW	~A

10. Invalid

```
*   1  ~☐A
    2  ☐(B ≡ A)
     [ ∴ ~◇B
*   3  asm: ◇B
*   4  ∴ ◇~A    {from 1}
    5  W ∴ B    {from 3}
    6  WW ∴ ~A    {from 4}
*   7  W ∴ (B ≡ A)    {from 2}
*   8  W ∴ (B ⊃ A)    {from 7}
    9  W ∴ (A ⊃ B)    {from 7}
   10  W ∴ A    {from 5 and 8}
*  11  WW ∴ (B ≡ A)    {from 2}
*  12  WW ∴ (B ⊃ A)    {from 11}
   13  WW ∴ (A ⊃ B)    {from 11}
   14  WW ∴ ~B    {from 6 and 12}
```

W	A, B
WW	~A, ~B

10.3b

1. Valid

```
*   1  (P ⊃ ☐(T ⊃ B))
*   2  ◇(T • ~B)
     [ ∴ ~P
    3  ┌ asm: P
    4  │ ∴ ☐(T ⊃ B)    {from 1 and 3}
*   5  │ W ∴ (T • ~B)    {from 2}
    6  │ W ∴ T    {from 5}
    7  │ W ∴ ~B    {from 5}
*   8  │ W ∴ (T ⊃ B)    {from 4}
    9  └ W ∴ B    {from 6 and 8}
```

10 ∴ ∼P {from 3; 7 contradicts 9}

3. Invalid

  1 ☐(B ⊃ B)
  [∴ (B ⊃ ☐B)
* 2 asm: ∼(B ⊃ ☐B)
  3 ∴ B {from 2}
* 4 ∴ ∼☐B {from 2}
* 5 ∴ ◇∼B {from 4}
  6 W ∴ ∼B {from 5}
  7 ∴ (B ⊃ B) {from 1}
  8 W ∴ (B ⊃ B) {from 1}

	B
W	∼B

5. Invalid

1 ☐(R ⊃ F)
2 ☐(U ⊃ ∼R)
[∴ ∼☐(F ⊃ U)
3 asm: ☐(F ⊃ U)
4 ∴ (R ⊃ F) {from 1}
5 ∴ (U ⊃ ∼R) {from 2}
6 ∴ (F ⊃ U) {from 3}
7 asm: ∼R {break 4}
8 asm: ∼F {break 6}

∼R, ∼F

10. Invalid

  1 ☐(D ∨ ∼D)
* 2 (☐D ⊃ ∼F)
* 3 (☐∼D ⊃ ∼F)
  [∴ ∼F
  4 asm: F
* 5 ∴ ∼☐D {from 2 and 4}
* 6 ∴ ∼☐∼D {from 2 and 4}
* 7 ∴ ◇∼D {from 3}
* 8 ∴ ◇D {from 4}
  9 W ∴ ∼D {from 5}
  10 WW ∴ D {from 6}
  11 W ∴ (D ∨ ∼D) {from 1}
  12 WW ∴ (D ∨ ∼D) {from 1}
  13 ∴ (D ∨ ∼D) {from 1}
  14 asm: D {break 13}

	F, D
W	∼D

WW | D

15. Valid

```
*  1   (M ⊃ ☐(S ⊃ G))
*  2   ◇(S • ~G)
       [ ∴ ~M
   3  ┌ asm: M
   4  │ ∴ ☐(S ⊃ G)    {from 1 and 3}
*  5  │ W ∴ (S • ~G)    {from 2}
   6  │ W ∴ S    {from 5}
   7  │ W ∴ ~G    {from 5}
*  8  │ W ∴ (S ⊃ G)    {from 4}
   9  └ W ∴ G    {from 6 and 8}
  10 ∴ ~M    {from 3; 7 contradicts 9}
```

20. Valid

```
*  1   (M ⊃ ☐(A ⊃ B))
*  2   ◇(~B • A)
       [ ∴ ~M
   3  ┌ asm: M
   4  │ ∴ ☐(A ⊃ B)    {from 1 and 3}
*  5  │ W ∴ (~B • A)    {from 2}
   6  │ W ∴ ~B    {from 5}
   7  │ W ∴ A    {from 5}
*  8  │ W ∴ (A ⊃ B)    {from 4}
   9  └ W ∴ ~A    {from 6 and 8}
  10 ∴ ~M    {from 3; 7 contradicts 9}
```

25. Valid

```
*  1   ◇(W • D)
   2   ☐(W ⊃ F) 0400
       [ ∴ ◇(F • D)
*  3  ┌ asm: ~◇(F • D)
*  4  │ W ∴ (W • D)    {from 1}
   5  │ ∴ ☐~(F • D)    {from 3}
   6  │ W ∴ W    {from 4}
   7  │ W ∴ D    {from 4}
*  8  │ W ∴ (W ⊃ F)    {from 2}
   9  │ W ∴ F    {from 6 and 8}
* 10  │ W ∴ ~(F • D)    {from 5}
  11  └ W ∴ ~F    {from 7 and 10}
  12 ∴ ◇(F • D)    {from 3; 9 contradicts 11}
```

Chapter 11 answers

11.1a

1. Valid in B or S5.

* 1 ◇☐A
[∴ A
2 ┌ asm: ~A
3 │ W ∴ ☐A {from 1} **# ⇒ W**
4 └ ∴ A {from 3} **Need B or S5**
5 ∴ A {from 2; 2 contradicts 4}

3. Valid in S4 or S5.

* 1 ◇◇A
[∴ ◇A
* 2 ┌ asm: ~◇A
* 3 │ W ∴ ◇A {from 1} **# ⇒ W**
4 │ ∴ ☐~A {from 2}
5 │ WW ∴ A {from 3} **W ⇒ WW**
6 └ WW ∴ ~A {from 4} **Need S4 or S5**
7 ∴ ◇A {from 2; 5 contradicts 6}

5. Valid in S5.

* 1 (☐A ⊃ ☐B)
[∴ ☐(☐A ⊃ ☐B)
* 2 ┌ asm: ~☐(☐A ⊃ ☐B)
* 3 │ ∴ ◇~(☐A ⊃ ☐B) {from 2}
* 4 │ W ∴ ~(☐A ⊃ ☐B) {from 3} **# ⇒ W**
5 │ W ∴ ☐A {from 4}
* 6 │ W ∴ ~☐B {from 4}
* 7 │ W ∴ ◇~B {from 6}
8 │ WW ∴ ~B {from 7} **W ⇒ WW**
9 │ ┌ asm: ~☐A {break 1}
10 │ │ ∴ ◇~A {from 9}
11 │ │ WWW ∴ ~A {from 10} **# ⇒ WWW**
12 │ └ WWW ∴ A {from 5} **Need S5**
13 │ ∴ ☐A {from 9; 11 contradicts 12}
14 │ ∴ ☐B {from 1 and 13}
15 └ WW ∴ B {from 14} **Need S4 or S5**
16 ∴ ☐(☐A ⊃ ☐B) {from 2; 8 contradicts 15}

10. Valid in B or S5.

* 1 ◇A
[∴ ◇☐◇A
* 2 ┌ asm: ~◇☐◇A
3 │ W ∴ A {from 1} **# ⇒ W**

```
   4 │ ∴ □~□◇A    {from 2}
*  5 │ W ∴ ~□◇A    {from 4}  Any system
*  6 │ W ∴ ◇~◇A    {from 5}
*  7 │ WW ∴ ~◇A    {from 6}  W ⇒ WW
   8 │ WW ∴ □~A    {from 7}
   9 └ W ∴ ~A    {from 8}  Need B or S5
  10 ∴ ◇□◇A    {from 2; 3 contradicts 9}
```

15. Valid in S4 or S5.

```
   1   □A
   [ ∴ □□□A
*  2 ┌ asm: ~□□□A
*  3 │ ∴ ◇~□□A    {from 2}
*  4 │ W ∴ ~□□A    {from 3}  # ⇒ W
*  5 │ W ∴ ◇~□A    {from 4}
*  6 │ WW ∴ ~□A    {from 5}  W ⇒ WW
*  7 │ WW ∴ ◇~A    {from 6}
   8 │ WWW ∴ ~A    {from 7}  WW ⇒ WWW
   9 └ WWW ∴ A    {from 1}  Need S4 or S5
  10 ∴ □□□A    {from 2; 8 contradicts 9}
```

11.1b

1. Valid in S5.

```
   1   □(N ⊃ □N)
*  2   ◇N
   [ ∴ □N
*  3 ┌ asm: ~□N
   4 │ W ∴ N    {from 2}  # ⇒ W
*  5 │ ∴ ◇~N    {from 3}
   6 │ WW ∴ ~N    {from 5}  # ⇒ WW
*  7 │ W ∴ (N ⊃ □N)    {from 1}  Any system
   8 │ W ∴ □N    {from 4 and 7}
   9 └ WW ∴ N    {from 8}  Need S5
  10 ∴ □N    {from 3; 6 contradicts 9}
```

3. This side is valid in S5.

```
   1   □(N ⊃ □N)
   [ ∴ ~(◇N • ◇~N)
*  2 ┌ asm: (◇N • ◇~N)
*  3 │ ∴ ◇N    {from 2}
*  4 │ ∴ ◇~N    {from 2}
   5 │ W ∴ N    {from 3}  # ⇒ W
   6 │ WW ∴ ~N    {from 4}  # ⇒ WW
*  7 │ W ∴ (N ⊃ □N)    {from 1}  Any system
   8 │ W ∴ □N    {from 5 and 7}
   9 └ WW ∴ N    {from 8}  Need S5
```

10 ∴ ~(◇N • ◇~N)   {from 2; 6 contradicts 9}

The other side is valid in S4 or S5.

* 1 ~(◇N • ◇~N)
[∴ ☐(N ⊃ ☐N)
* 2 ┌ asm: ~☐(N ⊃ ☐N)
* 3 │ ∴ ◇~(N ⊃ ☐N)   {from 2}
* 4 │ W ∴ ~(N ⊃ ☐N)   {from 3} **# ⇒ W**
5 │ W ∴ N   {from 4}
* 6 │ W ∴ ~☐N   {from 4} 0401
* 7 │ W ∴ ◇~N   {from 6}
8 │ WW ∴ ~N   {from 7} **W ⇒ WW**
9 │ ┌ asm: ~◇N   {break 1}
10 │ │ ∴ ☐~N   {from 9}
11 │ └ W ∴ ~N   {from 10} **Any system**
12 │ ∴ ◇N   {from 9; 5 contradicts 11}
* 13 │ ∴ ~◇~N   {from 1 and 12}
14 │ ∴ ☐N   {from 13}
15 └ WW ∴ N   {from 14} **Need S4 or S5**
16 ∴ ☐(N ⊃ ☐N)   {from 2; 8 contradicts 15}

5. Valid in S4 or S5.

1 ☐T
* 2 ◇~L
* 3 (N ⊃ ☐(☐T ⊃ L))
[∴ ~N
4 ┌ asm: N
5 │ W ∴ ~L   {from 2} **# ⇒ W**
6 │ ∴ ☐(☐T ⊃ L)   {from 3 and 4}
* 7 │ W ∴ (☐T ⊃ L)   {from 6} **Any system**
* 8 │ W ∴ ~☐T   {from 5 and 7}
* 9 │ W ∴ ◇~T   {from 8}
10 │ WW ∴ ~T   {from 9} **W ⇒ WW**
11 └ WW ∴ T   {from 1} **Need S4 or S5**
12 ∴ ~N   {from 4; 10 contradicts 11}

11.2a

1. (x)◇Ux
3. ☐Uj
5. (Ns • ◇~Ns)
10. (x)(Nx ⊃ ☐Ax)
15. ◇(x)(Cx ⊃ Tx)
20. (∃x)☐Ux

11.3a

1. Valid

```
* 1   (∃x)☐Fx
  [ ∴ ☐(∃x)Fx
* 2 ┌ asm: ~☐(∃x)Fx
  3 │ ∴ ☐Fa    {from 1}
* 4 │ ∴ ◇~(∃x)Fx    {from 2}
* 5 │ W ∴ ~(∃x)Fx    {from 4}
  6 │ W ∴ (x)~Fx    {from 5}
  7 │ W ∴ Fa    {from 3}
  8 └ W ∴ ~Fa    {from 6}
  9 ∴ ☐(∃x)Fx    {from 2; 7 contradicts 8}
```

3. Valid

```
  [ ∴ ☐(∃x)x=a
* 1 ┌ asm: ~☐(∃x)x=a
* 2 │ ∴ ◇~(∃x)x=a    {from 1}
* 3 │ W ∴ ~(∃x)x=a    {from 2}
  4 │ W ∴ (x)~x=a    {from 3}
  5 │ W ∴ ~a=a    {from 4}
  6 └ W ∴ a=a    {to contradict 5}
  7 ∴ ☐(∃x)x=a    {from 1; 5 contradicts 6}
```

5. Valid

```
* 1   ◇(x)Fx
  [ ∴ (x)◇Fx
* 2 ┌ asm: ~(x)◇Fx
  3 │ W ∴ (x)Fx    {from 1}
* 4 │ ∴ (∃x)~◇Fx    {from 2}
* 5 │ ∴ ~◇Fa    {from 4}
  6 │ ∴ ☐~Fa    {from 5}
  7 │ W ∴ Fa    {from 3}
  8 └ W ∴ ~Fa    {from 6}
  9 ∴ (x)◇Fx    {from 2; 7 contradicts 8}
```

10. Valid

```
* 1   (∃x)◇Fx
  [ ∴ ◇(∃x)Fx
* 2 ┌ asm: ~◇(∃x)Fx
* 3 │ ∴ ◇Fa    {from 1}
  4 │ ∴ ☐~(∃x)Fx    {from 2}
  5 │ W ∴ Fa    {from 3}
* 6 │ W ∴ ~(∃x)Fx    {from 4}
  7 │ W ∴ (x)~Fx    {from 6}
  8 └ W ∴ ~Fa    {from 7}
  9 ∴ ◇(∃x)Fx    {from 2; 5 contradicts 8}
```

11.3b

1. Invalid

1 Bi
[∴ ☐(x)(~Bx ⊃ ~x=i)
* 2 asm: ~☐(x)(~Bx ⊃ ~x=i)
* 3 ∴ ◇~(x)(~Bx ⊃ ~x=i) {from 2}
* 4 W ∴ ~(x)(~Bx ⊃ ~x=i) {from 3}
* 5 W ∴ (∃x)~(~Bx ⊃ ~x=i) {from 4}
* 6 W ∴ ~(~Ba ⊃ ~a=i) {from 5}
7 W ∴ ~Ba {from 6}
8 W ∴ a=i {from 6}
9 W ∴ ~(~Bi ⊃ ~i=i) {from 6 and 8}
10 W ∴ ~Bi {from 7 and 8}

	a, i
	Bi
W	~Bi, ~Ba a=i

3. Valid

1 ~a=p
* 2 (((∃x)☐x=p • ~(x)☐x=p) ⊃ S)
[∴ S
3 ┌ asm: ~S
* 4 │ ∴ ~((∃x)☐x=p • ~(x)☐x=p) {from 2 and 3}
5 │ ┌ asm: ~(∃x)☐x=p {break 4}
6 │ │ ∴ (x)~☐x=p {from 5}
7 │ │ ∴ ~☐p=p {from 6}
8 │ │ ∴ ◇~p=p {from 7}
9 │ │ W ∴ ~p=p {from 8}
10 │ └ W ∴ p=p {to contradict 9}
* 11 │ ∴ (∃x)☐x=p {from 5; 9 contradicts 10}
12 │ ∴ (x)☐x=p {from 4 and 11} 0402
13 │ ∴ ☐a=p {from 12}
14 └ ∴ a=p {from 13}
15 ∴ S {from 3; 1 contradicts 14}

5. Invalid

1 ☐(∃x)Ux
[∴ (∃x)☐Ux
* 2 asm: ~(∃x)☐Ux
3 ∴ (x)~☐Ux {from 3}
* 4 ∴ (∃x)Ux {from 1}
5 ∴ Ua {from 4}
* 6 ∴ ~☐Ua {from 3}
* 7 ∴ ◇~Ua {from 6}
8 W ∴ ~Ua {from 7}
* 9 W ∴ (∃x)Ux {from 1}

10 W ∴ Ub {from 1}

Endless loop: add "~Ub" to the actual world to make the conclusion false.

	a, b
	Ua, ~Ub
W	Ub, ~Ua

10. Valid

```
*  1   ◇(Ts • ~Bs)
   2   ☐(x)(Tx ⊃ Px)
   [ ∴ ~☐(x)(Px ⊃ Bx)
   3  ┌ asm: ☐(x)(Px ⊃ Bx)
*  4  │ W ∴ (Ts • ~Bs)   {from 1}
   5  │ W ∴ Ts   {from 4}
   6  │ W ∴ ~Bs   {from 4}
   7  │ W ∴ (x)(Tx ⊃ Px)   {from 2}
   8  │ W ∴ (x)(Px ⊃ Bx)   {from 3}
*  9  │ W ∴ (Ts ⊃ Ps)   {from 7}
  10  │ W ∴ Ps   {from 5 and 9}
* 11  │ W ∴ (Ps ⊃ Bs)   {from 8}
  12  └ W ∴ ~Ps   {from 6 and 11}
  13 ∴ ~☐(x)(Px ⊃ Bx)   {from 3; 10 contradicts 12}
```

15. Valid (but line 11 requires S5 or B).

```
*  1   ◇(∃x)Ux
   2   ☐(x)(Ux ⊃ ☐Ox)
   [ ∴ (∃x)Ox
*  3  ┌ asm: ~(∃x)Ox
*  4  │ W ∴ (∃x)Ux   {from 1}
   5  │ ∴ (x)~Ox   {from 3}
   6  │ W ∴ Ua   {from 4}
   7  │ W ∴ (x)(Ux ⊃ ☐Ox)   {from 2}
   8  │ ∴ ~Oa   {from 5}
*  9  │ W ∴ (Ua ⊃ ☐Oa)   {from 7}
  10  │ W ∴ ☐Oa   {from 6 and 9}
  11  └ ∴ Oa   {from 10}
  12∴(∃x)Ox   {from 3; 8 contradicts 11}
```

Chapter 12 answers

12.1a

1. (<u>**L**</u> ∨ <u>**S**</u>)
3. (<u>**A**</u> ⊃ W) *or, equivalently,* (~W ⊃ ~<u>**A**</u>)
5. ~(<u>**A**</u> • <u>**B**</u>)
10. ((x)Ax ⊃ A<u>**u**</u>)

15. (B ⊃ ~**A**)
20. (∃x)(S**x** • Wx)

12.2a

1. Valid

1 ~**A**
[∴ ~(**A** • **B**)
2 ┌ asm: (**A** • **B**)
3 └ ∴ **A** {from 2}
4 ∴ ~(**A** • **B**) {from 2; 1 contradicts 3}

3. Invalid

1 (A ⊃ **B**)
[∴ (~B ⊃ ~**A**)
* 2 asm: ~(~B ⊃ ~**A**)
3 ∴ ~B {from 2}
4 ∴ **A** {from 2}
5 asm: ~A {break 1}

~B, **A**, ~A

5. Valid

* 1 ~◇(**A** • **B**)
* 2 ~(**C** • ~**A**)
[∴ ~(**C** • **B**)
* 3 ┌ asm: (**C** • **B**)
4 │ ∴ □~(**A** • **B**) {from 1}
5 │ ∴ **C** {from 3}
6 │ ∴ **B** {from 3}
7 │ ∴ **A** {from 2 and 5}
* 8 │ ∴ ~(**A** • **B**) {from 4}
9 └ ∴ ~**A** {from 6 and 8}
10 ∴ ~(**C** • **B**) {from 3; 7 contradicts 9}

10. Valid

* 1 ~(**A** • ~**B**)
[∴ (~**A** ∨ **B**)
* 2 ┌ asm: ~(~**A** ∨ **B**)
3 │ ∴ **A** {from 2}
4 │ ∴ ~**B** {from 2}
5 └ ∴ **B** {from 1 and 3}
6 ∴ (~**A** ∨ **B**) {from 2; 4 contradicts 5}

12.2b

1. Valid

* 1 (**C** ∨ **E**)

2 V
* 3 (V ⊃ ~**C**)
[∴ **E**
4 ┌ asm: ~**E**
5 │ ∴ **C** {from 1 and 4}
6 └ ∴ ~**C** {from 2 and 3}
7 ∴ **E** {from 4; 5 contradicts 6}

3. Valid

* 1 (G ⊃ ~**E**) 0403
2 G
[∴ ~**E**
3 ┌ asm: **E**
4 └ ∴ ~**E** {from 1 and 2}
5 ∴ ~**E** {from 3; 3 contradicts 4}

5. Valid

* 1 (B ⊃ **C**)
2 B
[∴ **C**
3 ┌ asm: ~**C**
4 └ ∴ **C** {from 1 and 2}
5 ∴ **C** {from 3; 3 contradicts 4}

10. Valid

1 w=l
[∴ (G**u**w ∨ ~G**u**l)
* 2 ┌ asm: ~(G**u**w ∨ ~G**u**l)
3 │ ∴ ~G**u**w {from 2}
4 │ ∴ ~G**u**l {from 1 and 3}
5 └ ∴ G**u**l {from 2}
6 ∴ (G**u**w ∨ ~G**u**l) {from 2; 4 contradicts 5}

15. Invalid

1 (T ⊃ M)
2 **T**
[∴ **M**
3 asm: ~**M**
4 asm: ~T {break 1}

T, ~**M**, ~T

20. Invalid

1 (x)(Hx ⊃ E**x**)
[∴ (x)(~Ex ⊃ ~H**x**)
* 2 asm: ~(x)(~Ex ⊃ ~H**x**)
* 3 ∴ (∃x)~(~Ex ⊃ ~H**x**) {from 2}
* 4 ∴ ~(~Ea ⊃ ~H**a**) {from 3}
5 ∴ ~Ea {from 4}
6 ∴ H**a** {from 4}

7 ∴ (Ha ⊃ E$\underline{\mathbf{a}}$) {from 1}
8 asm: ~Ha {break 7}

a

~Ea, H$\underline{\mathbf{a}}$, ~Ha

12.3a

1. (A ⊃ O~$\underline{\mathbf{B}}$)
3. (O~$\underline{\mathbf{A}}$ ⊃ ~$\underline{\mathbf{A}}$)
5. ☐($\underline{\mathbf{A}}$ ⊃ R$\underline{\mathbf{A}}$)
10. O~($\underline{\mathbf{B}}$ • ~$\underline{\mathbf{A}}$)
15. (~◇(x)Ax ⊃ O~A$\underline{\mathbf{u}}$)
20. R(x)(~Tx ⊃ S$\underline{\mathbf{x}}$)

12.4a

1. Valid

1 O~$\underline{\mathbf{A}}$
[∴ O~($\underline{\mathbf{A}}$ • $\underline{\mathbf{B}}$)
* 2 asm: ~O~($\underline{\mathbf{A}}$ • $\underline{\mathbf{B}}$)
* 3 ∴ R($\underline{\mathbf{A}}$ • $\underline{\mathbf{B}}$) {from 2}
* 4 D ∴ ($\underline{\mathbf{A}}$ • $\underline{\mathbf{B}}$) {from 3}
5 D ∴ $\underline{\mathbf{A}}$ {from 4}
6 D ∴ $\underline{\mathbf{B}}$ {from 4}
7 D ∴ ~$\underline{\mathbf{A}}$ {from 1}
8 ∴ O~($\underline{\mathbf{A}}$ • $\underline{\mathbf{B}}$) {from 2; 5 contradicts 7}

3. Valid

1 b=c
[∴ (OF$\underline{\mathbf{a}}$b ⊃ OF$\underline{\mathbf{a}}$c)
* 2 asm: ~(OF$\underline{\mathbf{a}}$b ⊃ OF$\underline{\mathbf{a}}$c)
3 ∴ OF$\underline{\mathbf{a}}$b {from 2}
4 ∴ ~OF$\underline{\mathbf{a}}$c {from 2}
5 ∴ OF$\underline{\mathbf{a}}$c {from 1 and 3}
6 ∴ (OF$\underline{\mathbf{a}}$b ⊃ OF$\underline{\mathbf{a}}$c) {from 2; 4 contradicts 5}

5. Invalid

[∴ O($\underline{\mathbf{A}}$ ⊃ O$\underline{\mathbf{A}}$)
* 1 asm: ~O($\underline{\mathbf{A}}$ ⊃ O$\underline{\mathbf{A}}$)
* 2 ∴ R~($\underline{\mathbf{A}}$ ⊃ O$\underline{\mathbf{A}}$) {from 1}
* 3 D ∴ ~($\underline{\mathbf{A}}$ ⊃ O$\underline{\mathbf{A}}$) {from 2}
4 D ∴ $\underline{\mathbf{A}}$ {from 3}
* 5 D ∴ ~O$\underline{\mathbf{A}}$ {from 3}
* 6 D ∴ R~$\underline{\mathbf{A}}$ {from 5}
7 DD ∴ ~$\underline{\mathbf{A}}$ {from 6}

10. Valid

* 1 (A ⊃ O$\underline{\mathbf{B}}$)

[∴ O(A ⊃ **B**)
* 2 ┌ asm: ~O(A ⊃ **B**)
* 3 │ ∴ R~(A ⊃ **B**) {from 2}
* 4 │ D ∴ ~(A ⊃ **B**) {from 3}
5 │ D ∴ A {from 4}
6 │ D ∴ ~**B** {from 4}
7 │ ∴ A {from 5 by indicative transfer}
8 │ ∴ O**B** {from 1 and 7}
9 └ D ∴ **B** {from 8}
10 ∴ O(A ⊃ **B**) {from 2; 6 contradicts 9}

15. Valid

1 O**A**
2 O**B**
[∴ ◇(A • B)
3 ┌ asm: ~◇(A • B)
4 │ ┌ asm: O(**A** • **B**) {assume to get opposite using 3 and Kant's Law}
5 │ └ ◇(A • B) {from 4 using Kant's Law}
* 6 │ ∴ ~O(**A** • **B**) {from 4; 3 contradicts 5}
* 7 │ ∴ R~(**A** • **B**) {from 6}
* 8 │ D ∴ ~(**A** • **B**) {from 7}
9 │ D ∴ **A** {from 1}
10 │ D ∴ **B** {from 2}
11 └ D ∴ ~**B** {from 8 and 9}
12 ∴ ◇(A • B) {from 3; 10 contradicts 11}

20. Invalid

1 O(x)(Fx ⊃ G**x**) 0404
2 OF**a**
[∴ OG**a**
* 3 asm: ~OG**a**
* 4 ∴ R~G**a** {from 3}
5 D ∴ ~G**a** {from 4}
6 D ∴ (x)(Fx ⊃ G**x**) {from 1}
7 D ∴ F**a** {from 2}
* 8 D ∴ (Fa ⊃ G**a**) {from 6}
9 D ∴ ~Fa {from 5 and 8}
10 ∴ ~Fa {from 9 by indicative transfer}

25. Valid

* 1 (A ∨ O**B**)
2 ~A
[∴ O**B**
* 3 ┌ asm: ~O**B**
4 └ ∴ O**B** {from 1 and 2}
5 ∴ O**B** {from 3; 3 contradicts 4}

12.4b

1. Valid

* 1 ~R(**T** • **D**)
2 O**D**
[∴ ~**T**
3 asm: **T**
4 ∴ O~(**T** • **D**) {from 1}
5 ∴ **D** {from 2}
* 6 ∴ ~(**T** • **D**) {from 4}
7 ∴ ~**D** {from 3 and 6}
8 ∴ ~**T** {from 3; 5 contradicts 7}

3. Valid

1 A
2 O~**A**
* 3 ((A • ◇~A) ⊃ F)
[∴ F
4 asm: ~F
* 5 ∴ ~(A • ◇~A) {from 3 and 4}
* 6 ∴ ~◇~A {from 1 and 5}
7 ∴ ◇~A {from 2 by Kant's Law}
8 ∴ F {from 4; 6 contradicts 7}

5. Invalid

[∴ (O**A** ⊃ A)
* 1 asm: ~(O**A** ⊃ A)
2 ∴ O**A** {from 1}
3 ∴ ~A {from 1}
4 ∴ **A** {from 2}

10. Invalid

* 1 R(∃x)A**x**
[∴ (x)RA**x**
* 2 asm: ~(x)RA**x**
* 3 D ∴ (∃x)A**x** {from 1}
* 4 ∴ (∃x)~RA**x** {from 2}
5 D ∴ A**a** {from 3}
* 6 ∴ ~RA**b** {from 4}
7 ∴ O~A**b** {from 6}
8 D ∴ ~A**b** {from 7}

15. Valid

* 1 (O~**K** ⊃ O~**I**)
* 2 (R**N** ⊃ R**I**)
[∴ (O~**K** ⊃ O~**N**)
* 3 asm: ~(O~**K** ⊃ O~**N**)
4 ∴ O~**K** {from 3}
* 5 ∴ ~O~**N** {from 3}

* 6 | ∴ R**N** {from 5}
7 | ∴ O~**I** {from 1 and 4}
* 8 | ∴ R**I** {from 2 and 6}
9 | DD ∴ **I** {from 8}
10 └ DD ∴ ~**I** {from 7}
11 ∴ (O~**K** ⊃ O~**N**) {from 3; 9 contradicts 10}

20. Valid

1 N
2 ☐(**B** ⊃ **D**)
3 ☐(**D** ⊃ (N ⊃ **S**))
[∴ O(**S** ∨ ~**B**)
* 4 ┌ asm: ~O(**S** ∨ ~**B**)
* 5 | ∴ R~(**S** ∨ ~**B**) {from 4}
* 6 | D ∴ ~(**S** ∨ ~**B**) {from 5}
7 | D ∴ ~**S** {from 6}
8 | D ∴ **B** {from 6}
* 9 | D ∴ (**B** ⊃ **D**) {from 2}
10 | D ∴ **D** {from 8 and 9}
* 11 | D ∴ (**D** ⊃ (N ⊃ **S**)) {from 3}
* 12 | D ∴ (N ⊃ **S**) {from 10 and 11}
13 | D ∴ ~N {from 7 and 12}
14 └ ∴ ~N {from 13 by indicative transfer}
15 ∴ O(**S** ∨ ~**B**) {from 4; 1 contradicts 14}

25. Valid

* 1 (R**P** ⊃ O**S**)
* 2 ~◇S
[∴ O~**P**
* 3 ┌ asm: ~O~**P**
4 | ∴ ☐~S {from 2}
* 5 | ∴ R**P** {from 3}
6 | ∴ O**S** {from 1 and 5}
* 7 | ∴ ◇S {from 6 by Kant's Law}
8 | W ∴ S {from 7}
9 └ W ∴ ~S {from 4}
10 ∴ O~**P** {from 3; 8 contradicts 9}

Chapter 13 answers

13.1a

1. u:~G
3. ~u:G
5. ☐(u:G ⊃ ~u:~G)
10. ~(**u**:A • **u**:~A) 0405

13.2a

1. Valid

```
*  1   ~◇(A • B)
       [ ∴ ~(u:A • u:B)
*  2  ┌ asm: (u:A • u:B)
   3  │ ∴ □~(A • B)    {from 1}
   4  │ ∴ u:A    {from 2}
   5  │ ∴ u:B    {from 2}
   6  │ u ∴ A    {from 4}
*  7  │ u ∴ ~(A • B)    {from 3}
   8  │ u ∴ ~B    {from 6 and 7}
   9  └ u ∴ B    {from 5}
  10 ∴ ~(u:A • u:B)    {from 2; 8 contradicts 9}
```

3. Invalid

```
* 1   ~◇(A • B)
      [ ∴ (u:A ⊃ ~u:B)
* 2   asm: ~(u:A ⊃ ~u:B)
  3   ∴ □~(A • B)    {from 1}
  4   ∴ u:A    {from 2}
  5   ∴ u:B    {from 2}
  6   u ∴ B    {from 5}
* 7   u ∴ ~(A • B)    {from 3}
  8   u ∴ ~A    {from 6 and 7}
```

5. Invalid

```
*  1   ~◇(A • B)
       [ ∴ (u:~A ∨ u:~B)
*  2   asm: ~(u:~A ∨ u:~B)
   3   ∴ □~(A • B)    {from 1}
*  4   ∴ ~u:~A    {from 2}
*  5   ∴ ~u:~B    {from 2}
   6   u ∴ A    {from 4}
   7   uu ∴ B    {from 5}
*  8   u ∴ ~(A • B)    {from 3}
   9   u ∴ ~B    {from 6 and 8}
* 10   uu ∴ ~(A • B)    {from 3}
  11   uu ∴ ~A    {from 7 and 10}
```

10. Valid

```
*  1   ~◇(A • B)
       [ ∴ ~(u:A • ~u:~B)
*  2  ┌ asm: (u:A • ~u:~B)
   3  │ ∴ □~(A • B)    {from 1}
   4  │ ∴ u:A    {from 2}
*  5  │ ∴ ~u:~B    {from 2}
   6  │ u ∴ B    {from 5}
*  7  │ u ∴ ~(A • B)    {from 3}
   8  │ u ∴ ~A    {from 6 and 7}
```

9 └ u ∴ A {from 4}
10 ∴ ~(**u**:A • ~**u**:~B) {from 2; 8 contradicts 9}

13.2b

1. Valid

1 ☐(A ⊃ B)
* 2 ~**u**:B
[∴ ~**u**:A
3 ┌ asm: **u**:A
4 │ u ∴ ~B {from 2}
* 5 │ u ∴ (A ⊃ B) {from 1}
6 │ u ∴ ~A {from 4 and 5}
7 └ u ∴ A {from 3}
8 ∴ ~**u**:A {from 3; 6 contradicts 7}

3. Invalid

1 u:A
[∴ ~**u**:~A
2 asm: **u**:~A
3 u ∴ ~A {from 2}

5. Invalid

[∴ (**u**:A ∨ **u**:~A)
* 1 asm: ~(**u**:A ∨ **u**:~A)
* 2 ∴ ~**u**:A {from 1}
* 3 ∴ ~**u**:~A {from 1}
4 u ∴ ~A {from 2}
5 uu ∴ A {from 3}

10. Invalid

[∴ (A ⊃ **u**:A)
* 1 asm: ~(A ⊃ **u**:A)
2 ∴ A {from 1}
* 3 ∴ ~**u**:A {from 1}
4 u ∴ ~A {from 3}

13.3a

1. u:S**a**
3. u:OS**a**
5. **u**:S**a**
10. (u:OA**u** ⊃ A**u**)
15. (u:A**x**u ⊃ A**u**x)

13.4a

1. Valid

[∴ ~($\underline{\mathbf{u}}$:$\underline{\mathbf{A}}$ • $\underline{\mathbf{u}}$:~$\underline{\mathbf{A}}$)
* 1 asm: ($\underline{\mathbf{u}}$:$\underline{\mathbf{A}}$ • $\underline{\mathbf{u}}$:~$\underline{\mathbf{A}}$)
2 ∴ $\underline{\mathbf{u}}$:$\underline{\mathbf{A}}$ {from 1}
3 ∴ $\underline{\mathbf{u}}$:~$\underline{\mathbf{A}}$ {from 1}
4 u ∴ $\underline{\mathbf{A}}$ {from 2}
5 u ∴ ~$\underline{\mathbf{A}}$ {from 3}
6 ∴ ~($\underline{\mathbf{u}}$:$\underline{\mathbf{A}}$ • $\underline{\mathbf{u}}$:~$\underline{\mathbf{A}}$) {from 1; 4 contradicts 5} 0406

3. Invalid

[∴ ($\underline{\mathbf{u}}$:B$\underline{\mathbf{a}}$ ∨ $\underline{\mathbf{u}}$:~B$\underline{\mathbf{a}}$)
* 1 asm: ~($\underline{\mathbf{u}}$:B$\underline{\mathbf{a}}$ ∨ $\underline{\mathbf{u}}$:~B$\underline{\mathbf{a}}$)
* 2 ∴ ~$\underline{\mathbf{u}}$:B$\underline{\mathbf{a}}$ {from 1}
* 3 ∴ ~$\underline{\mathbf{u}}$:~B$\underline{\mathbf{a}}$ {from 1}
4 u ∴ ~B$\underline{\mathbf{a}}$ {from 2}
5 uu ∴ B$\underline{\mathbf{a}}$ {from 3}

5. Invalid

1 u:(x)OA$\underline{\mathbf{x}}$
[∴ $\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$
* 2 asm: ~$\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$
3 u ∴ ~A$\underline{\mathbf{u}}$ {from 2}

10. Valid

1 ☐($\underline{\mathbf{A}}$ ⊃ $\underline{\mathbf{B}}$)
[∴ ~($\underline{\mathbf{u}}$:O$\underline{\mathbf{A}}$ • ~$\underline{\mathbf{u}}$:$\underline{\mathbf{B}}$)
* 2 asm: ($\underline{\mathbf{u}}$:O$\underline{\mathbf{A}}$ • ~$\underline{\mathbf{u}}$:$\underline{\mathbf{B}}$)
3 ∴ $\underline{\mathbf{u}}$:O$\underline{\mathbf{A}}$ {from 2}
* 4 ∴ ~$\underline{\mathbf{u}}$:$\underline{\mathbf{B}}$ {from 2}
5 u ∴ ~$\underline{\mathbf{B}}$ {from 4}
* 6 u ∴ ($\underline{\mathbf{A}}$ ⊃ $\underline{\mathbf{B}}$) {from 1}
7 u ∴ ~$\underline{\mathbf{A}}$ {from 5 and 6}
8 u ∴ O$\underline{\mathbf{A}}$ {from 3}
9 u ∴ $\underline{\mathbf{A}}$ {from 8}
10 ∴ ~($\underline{\mathbf{u}}$:O$\underline{\mathbf{A}}$ • ~$\underline{\mathbf{u}}$:$\underline{\mathbf{B}}$) {from 2; 7 contradicts 9}

13.4b

1. Valid

[∴ ~($\underline{\mathbf{u}}$:(x)OA$\underline{\mathbf{x}}$ • ~$\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$)
* 1 asm: ($\underline{\mathbf{u}}$:(x)OA$\underline{\mathbf{x}}$ • ~$\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$)
2 ∴ $\underline{\mathbf{u}}$:(x)OA$\underline{\mathbf{x}}$ {from 1}
* 3 ∴ ~$\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$ {from 1}
4 u ∴ ~A$\underline{\mathbf{u}}$ {from 3}
5 u ∴ (x)OA$\underline{\mathbf{x}}$ {from 2}
6 u ∴ OA$\underline{\mathbf{u}}$ {from 5}
7 u ∴ A$\underline{\mathbf{u}}$ {from 6}

8 ∴ ~(**u**:(x)OA**x** • ~**u**:A**u**) {from 1; 4 contradicts 7}

3. Valid

1 ☐(**E** ⊃ (N ⊃ **M**))
[∴ ~((**u**:**E** • **u**:N) • ~**u**:**M**)
* 2 asm: ((**u**:**E** • **u**:N) • ~**u**:**M**)
* 3 ∴ (**u**:**E** • **u**:N) {from 2}
* 4 ∴ ~**u**:**M** {from 2}
5 ∴ **u**:**E** {from 3}
6 ∴ **u**:N {from 3}
7 u ∴ ~**M** {from 4}
* 8 u ∴ (**E** ⊃ (N ⊃ **M**)) {from 1}
9 u ∴ **E** {from 5}
* 10 u ∴ (N ⊃ **M**) {from 8 and 9}
11 u ∴ ~N {from 7 and 10}
12 u ∴ N {from 6}
13 ∴ ~((**u**:**E** • **u**:N) • ~**u**:**M**) {from 2; 11 contradicts 12}

5. Valid

[∴ ~(**u**:(x)O~K**x** • ~**u**:(N ⊃ ~K**u**))
* 1 asm: (**u**:(x)O~K**x** • ~**u**:(N ⊃ ~K**u**))
2 ∴ **u**:(x)O~K**x** {from 1}
* 3 ∴ ~**u**:(N ⊃ ~K**u**) {from 1}
* 4 u ∴ ~(N ⊃ ~K**u**) {from 3}
5 u ∴ N {from 4}
6 u ∴ K**u** {from 4}
7 u ∴ (x)O~K**x** {from 2}
8 u ∴ O~K**u** {from 7}
9 u ∴ ~K**u** {from 8}
10 ∴ ~(**u**:(x)O~K**x** • ~**u**:(N ⊃ ~K**u**)) {from 1; 6 contradicts 9}

10. Invalid

[∴ (u:A**u** ⊃ **u**:RA**u**)
* 1 asm: ~(u:A**u** ⊃ **u**:RA**u**)
2 ∴ u:A**u** {from 1}
* 3 ∴ ~**u**:RA**u** {from 1}
* 4 u ∴ ~RA**u** {from 3}
5 u ∴ O~A**u** {from 4}

13.5a

1. O**u**:Sa
3. R**u**:OS**a**
5. (x)~R**x**:G
10. (O**u**:x=x • (x=x • u:x=x))
15. (O**u**:A ≡ ~◇(u:A • ~A))
20. (u:A**x**u ⊃ OA**u**x)
25. ((~Du • O**u**:Bj) ⊃ O**u**:Fj)

13.6a

1. Valid

1 ☐(A ⊃ B)
* 2 ~R$\underline{\mathbf{u}}$:B
[∴ ~R$\underline{\mathbf{u}}$:A
* 3 asm: R$\underline{\mathbf{u}}$:A
4 ∴ O~$\underline{\mathbf{u}}$:B {from 2}
5 D ∴ $\underline{\mathbf{u}}$:A {from 3}
* 6 D ∴ ~$\underline{\mathbf{u}}$:B {from 4}
7 Du ∴ ~B {from 6}
* 8 Du ∴ (A ⊃ B) {from 1}
9 Du ∴ ~A {from 7 and 8}
10 Du ∴ A {from 5}
11 ∴ ~R$\underline{\mathbf{u}}$:A {from 3; 9 contradicts 10}

3. Valid

* 1 R(~$\underline{\mathbf{u}}$:A • ~$\underline{\mathbf{u}}$:~A)
[∴ ~O$\underline{\mathbf{u}}$:A
2 asm: O$\underline{\mathbf{u}}$:A
3 D ∴ (~$\underline{\mathbf{u}}$:A • ~$\underline{\mathbf{u}}$:~A) {from 1}
4 D ∴ $\underline{\mathbf{u}}$:A {from 2} 0407
5 D ∴ ~$\underline{\mathbf{u}}$:A {from 3}
6 ∴ ~O$\underline{\mathbf{u}}$:A {from 2; 4 contradicts 5}

5. Invalid

1 O$\underline{\mathbf{a}}$:(C • D)
[∴ O$\underline{\mathbf{b}}$:C
* 2 asm: ~O$\underline{\mathbf{b}}$:C
* 3 ∴ R~$\underline{\mathbf{b}}$:C {from 2}
* 4 D ∴ ~$\underline{\mathbf{b}}$:C {from 3}
5 Db ∴ ~C {from 4}
6 D ∴ $\underline{\mathbf{a}}$:(C • D) {from 1}
* 7 Da ∴ (C • D) {from 6}
8 Da ∴ C {from 7}
9 Da ∴ D {from 7}

10. Valid

1 O$\underline{\mathbf{u}}$:(A ⊃ OB$\underline{\mathbf{u}}$)
[∴ ~($\underline{\mathbf{u}}$:A • ~$\underline{\mathbf{u}}$:B$\underline{\mathbf{u}}$)
* 2 asm: ($\underline{\mathbf{u}}$:A • ~$\underline{\mathbf{u}}$:B$\underline{\mathbf{u}}$)
3 ∴ $\underline{\mathbf{u}}$:A {from 2}
* 4 ∴ ~$\underline{\mathbf{u}}$:B$\underline{\mathbf{u}}$ {from 2}
5 u ∴ ~B$\underline{\mathbf{u}}$ {from 4}
6 ∴ $\underline{\mathbf{u}}$:(A ⊃ OB$\underline{\mathbf{u}}$) {from 1}
7 u ∴ A {from 3}
* 8 u ∴ (A ⊃ OB$\underline{\mathbf{u}}$) {from 6}
9 u ∴ OB$\underline{\mathbf{u}}$ {from 7 and 8}
10 u ∴ B$\underline{\mathbf{u}}$ {from 9}
11 ∴ ~($\underline{\mathbf{u}}$:A • ~$\underline{\mathbf{u}}$:B$\underline{\mathbf{u}}$) {from 2; 5 contradicts 10}

13.6b

1. Valid

1 Ou:G
[∴ ~Ru:~G
* 2 ┌ asm: Ru:~G
3 │ D ∴ u:~G {from 2}
4 │ D ∴ u:G {from 1}
5 │ Du ∴ ~G {from 3}
6 └ Du ∴ G {from 4}
7 ∴ ~Ru:~G {from 2; 5 contradicts 6}

3. Valid

[∴ O~(u:OAu • ~u:Au)
* 1 ┌ asm: ~O~(u:OAu • ~u:Au)
* 2 │ ∴ R(u:OAu • ~u:Au) {from 1}
* 3 │ D ∴ (u:OAu • ~u:Au) {from 2}
4 │ D ∴ u:OAu {from 3}
* 5 │ D ∴ ~u:Au {from 3}
6 │ Du ∴ ~Au {from 5}
7 │ Du ∴ OAu {from 4}
8 └ Du ∴ Au {from 7}
9 ∴ O~(u:OAu • ~u:Au) {from 1; 6 contradicts 8}

5. Valid

1 ☐(E ⊃ C)
* 2 Ou:E
[∴ Ou:C
* 3 ┌ asm: ~Ou:C
* 4 │ ∴ R~u:C {from 3}
* 5 │ D ∴ ~u:C {from 4}
6 │ Du ∴ ~C {from 5}
* 7 │ Du ∴ (E ⊃ C) {from 1}
8 │ Du ∴ ~E {from 6 and 7}
9 │ D ∴ u:E {from 2}
10 └ Du ∴ E {from 9}
11 ∴ ~Ou:E {from 3; 8 contradicts 10}

10. Valid

1 a:F
2 Oa:F
3 ~F
4 C
5 a:(F ∨ C)
6 ~K
[∴ (((Oa:(F ∨ C) • (F ∨ C)) • a:(F ∨ C)) • ~K)
* 7 ┌ asm: ~(((Oa:(F ∨ C) • (F ∨ C)) • a:(F ∨ C)) • ~K)
* 8 │ ∴ ~((Oa:(F ∨ C) • (F ∨ C)) • a:(F ∨ C)) {from 6 and 7}
* 9 │ ∴ ~(Oa:(F ∨ C) • (F ∨ C)) {from 5 and 8}
10 │ ┌ asm: ~(F ∨ C) {break 9}
11 │ └ ∴ ~C {from 10}
12 │ ∴ (F ∨ C) {from 10; 4 contradicts 11}

* 13 │ ∴ ~O$\underline{a}$:(F ∨ C)   {from 9 and 12}
* 14 │ ∴ R~$\underline{a}$:(F ∨ C)   {from 13}
* 15 │ D ∴ ~$\underline{a}$:(F ∨ C)   {from 14}
16 │ D ∴ $\underline{a}$:F   {from 2}
17 │ Du ∴ ~(F ∨ C)   {from 15}
18 │ Du ∴ F   {from 16}
19 └ Du ∴ ~F   {from 17}
20 ∴ (((O$\underline{a}$:(F ∨ C) • (F ∨ C)) • a:(F ∨ C)) • ~K)   {from 7; 18 contradicts 19}

15. Valid

1 O$\underline{u}$:(M ⊃ F)
[∴ ~($\underline{u}$:M • ~$\underline{u}$:F)
* 2 ┌ asm: ($\underline{u}$:M • ~$\underline{u}$:F)
3 │ ∴ $\underline{u}$:M   {from 2}
* 4 │ ∴ ~$\underline{u}$:F   {from 2}
5 │ u ∴ ~F   {from 4}
6 │ ∴ $\underline{u}$:(M ⊃ F)   {from 1}
7 │ u ∴ M   {from 3}
* 8 │ u ∴ (M ⊃ F)   {from 6}
9 └ u ∴ ~M   {from 5 and 8}
10 ∴ ~($\underline{u}$:M • ~$\underline{u}$:F)   {from 2; 7 contradicts 9} 0408

20. Invalid

1 O~($\underline{u}$:A • ~$\underline{u}$:B)
[∴ (u:A ⊃ u:B)
* 2 asm: ~(u:A ⊃ u:B)
3 ∴ u:A   {from 2}
4 ∴ ~u:B   {from 2}
5 ∴ ~($\underline{u}$:A • ~$\underline{u}$:B)   {from 1}
** 6 asm: ~$\underline{u}$:A   {break 5}
7 u ∴ ~A   {from 6}

25. Valid

1 O$\underline{u}$:(P ⊃ OA$\underline{u}$)
2 O$\underline{u}$:P
[∴ $\underline{u}$:A$\underline{u}$
* 3 ┌ asm: ~$\underline{u}$:A$\underline{u}$
4 │ u ∴ ~A$\underline{u}$ {from 3}
5 │ ∴ $\underline{u}$:(P ⊃ OA$\underline{u}$)   {from 1}
6 │ ∴ $\underline{u}$:P   {from 2}
* 7 │ u ∴ (P ⊃ OA$\underline{u}$)   {from 5}
8 │ u ∴ P   {from 6}
9 │ u ∴ OA$\underline{u}$ {from 7 and 8}
10 └ u ∴ A$\underline{u}$ {from 9}
11 ∴ $\underline{u}$:A$\underline{u}$ {from 3; 4 contradicts 10}

30. Invalid

1 (x)R$\underline{x}$:A
[∴ R(x)$\underline{x}$:A
* 2 asm: ~R(x)$\underline{x}$:A
3 ∴ O~(x)$\underline{x}$:A   {from 2}
* 4 ∴ R$\underline{a}$:A   {from 1}

5 D ∴ $\underline{a}$:A {from 4}
* 6 D ∴ ~(x)$\underline{x}$:A {from 3}
* 7 D ∴ (∃x)~$\underline{x}$:A {from 6}
* 8 D ∴ ~$\underline{b}$:A {from 7}
9 Db ∴ ~A {from 8}
* 10 ∴ R$\underline{b}$:A {from 1}
11 DD ∴ $\underline{b}$:A {from 10}

Endless loop: add "~$\underline{a}$:A" to world DD to make the conclusion false. (You weren't required to give a refutation.)

a, b

D	$\underline{a}$:A, ~$\underline{b}$:A
DD	$\underline{b}$:A, ~$\underline{a}$:A

Chapter 14 answers

14.6

Impartiality formula – valid. (See footnote at the end of Chapter 14.)

[∴ ~($\underline{u}$:R$\underline{A}$ • ~$\underline{u}$:(∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)))
* 1 asm: ($\underline{u}$:R$\underline{A}$ • ~$\underline{u}$:(∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)))
2 ∴ $\underline{u}$:R$\underline{A}$ {from 1}
* 3 ∴ ~$\underline{u}$:(∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)) {from 1}
* 4 u ∴ ~(∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)) {from 3}
5 u ∴ R$\underline{A}$ {from 2}
6 u ∴ (∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)) {from 5 by G5}
7 ∴ ~($\underline{u}$:R$\underline{A}$ • ~$\underline{u}$:(∃F)(F$\underline{A}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$))) {from 1; 4 contradicts 6}

Formula of universal law – valid. (See footnote at the end of Chapter 14.)

[∴ ~($\underline{u}$:A$\underline{u}$ • ~$\underline{u}$:(∃F)(F*A$\underline{u}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ M$\underline{X}$)))
* 1 asm: ($\underline{u}$:A$\underline{u}$ • ~$\underline{u}$:(∃F)(F*A$\underline{u}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ M$\underline{X}$)))
2 ∴ $\underline{u}$:A$\underline{u}$ {from 1}
* 3 ∴ ~$\underline{u}$:(∃F)(F*A$\underline{u}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ M$\underline{X}$)) {from 1}
* 4 u ∴ ~(∃F)(F*A$\underline{u}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ M$\underline{X}$)) {from 3}
5 u ∴ A$\underline{u}$ {from 2}
6 u asm: ~RA$\underline{u}$
7 u ∴ O~A$\underline{u}$ {from 6}
8 u ∴ ~A$\underline{u}$ {from 7}
9 u ∴ RA$\underline{u}$ {from 6; 5 contradicts 8}
* 10 u ∴ (∃F)(FA$\underline{u}$ • ■($\underline{X}$)(F$\underline{X}$ ⊃ R$\underline{X}$)) {from 9 by G5}
* 11 u ∴ (GA$\underline{u}$ • ■($\underline{X}$)(G$\underline{X}$ ⊃ R$\underline{X}$)) {from 10}
12 u ∴ GA$\underline{u}$ {from 11}
13 u ∴ ■($\underline{X}$)(G$\underline{X}$ ⊃ R$\underline{X}$) {from 11}
14 u ∴ ($\underline{X}$)(∃F)F*$\underline{X}$ {rule G11}
* 15 u ∴ (∃F)F*A$\underline{u}$ {from 14}
16 u ∴ H*A$\underline{u}$ {from 15}
* 17 u ∴ (HA$\underline{u}$ • (F)(FA$\underline{u}$ ⊃ □($\underline{X}$)(H$\underline{X}$ ⊃ F$\underline{X}$))) {from 16 by G10}
18 u ∴ HA$\underline{u}$ {from 17}
19 u ∴ (F)(FA$\underline{u}$ ⊃ □($\underline{X}$)(H$\underline{X}$ ⊃ F$\underline{X}$)) {from 17}
* 20 u ∴ (GA$\underline{u}$ ⊃ □($\underline{X}$)(H$\underline{X}$ ⊃ G$\underline{X}$)) {from 19}
21 u ∴ □($\underline{X}$)(H$\underline{X}$ ⊃ G$\underline{X}$) {from 12 and 20}

22 u ∴ (F)~(F*A$\underline{\mathbf{u}}$ • ■($\underline{\mathbf{X}}$)(F$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$)) {from 4}

* 23 u ∴ ~(H*A$\underline{\mathbf{u}}$ • ■($\underline{\mathbf{X}}$)(H$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$)) {fm 22}

24 u ∴ ~■($\underline{\mathbf{X}}$)(H$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$) {fm 16 & 23}

* 25 uH ∴ ~($\underline{\mathbf{X}}$)(H$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$) {fm 24 by G8}

* 26 uH ∴ (∃$\underline{\mathbf{X}}$)~(H$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$) {from 25}

* 27 uH ∴ ~(H$\underline{\mathbf{B}}$ ⊃ M$\underline{\mathbf{B}}$) {from 26}

28 uH ∴ H$\underline{\mathbf{B}}$ {from 27}

29 uH ∴ ~M$\underline{\mathbf{B}}$ {from 27}

30 uH ∴ ($\underline{\mathbf{X}}$)(H$\underline{\mathbf{X}}$ ⊃ G$\underline{\mathbf{X}}$) {from 21}

* 31 uH ∴ (H$\underline{\mathbf{B}}$ ⊃ G$\underline{\mathbf{B}}$) {from 30}

32 uH ∴ G$\underline{\mathbf{B}}$ {from 28 and 31}

33 uH ∴ ($\underline{\mathbf{X}}$)(G$\underline{\mathbf{X}}$ ⊃ R$\underline{\mathbf{X}}$) {from 13 by G7}

* 34 uH ∴ (G$\underline{\mathbf{B}}$ ⊃ R$\underline{\mathbf{B}}$) {from 33}

35 uH ∴ R$\underline{\mathbf{B}}$ {from 32 and 34}

36 uH ∴ M$\underline{\mathbf{B}}$ {from 35 by G1}

37 ∴ ~($\underline{\mathbf{u}}$:A$\underline{\mathbf{u}}$ • ~$\underline{\mathbf{u}}$:(∃F)(F*A$\underline{\mathbf{u}}$ • ■($\underline{\mathbf{X}}$)(F$\underline{\mathbf{X}}$ ⊃ M$\underline{\mathbf{X}}$))) {from 1; 29 contradicts 36}

Index