

Counterfactuals

David Lewis

B

Counterfactuals

Counterfactuals

DAVID LEWIS



© 1973 by David Lewis

BLACKWELL PUBLISHING

350 Main Street, Malden, MA 02148-5020, USA

108 Cowley Road, Oxford OX4 1JF, UK

550 Swanston Street, Carlton, Victoria 3053, Australia

The right of David Lewis to be identified as the Author of this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

First published by Basil Blackwell Ltd 1973

First publishing in the USA by Harvard University Press 1973

Reissued by Blackwell Publishers 2001

Reprinted 2003, 2005

Library of Congress Cataloging-in-Publication Data

Lewis, David K., 1941-

Counterfactuals / David Lewis.

p. cm.

First published: 1973.

Includes bibliographical references and index.

ISBN 0-631-22495-5 (hardback: alk. paper) — ISBN 0-631-22425-4 (pbk. : alk. paper)

1. Counterfactuals (Logic) I. Title.

BC199.C66 L48 2000

160—dc21

00-059899

A catalogue record for this title is available from the British Library.

Printed and bound in the United Kingdom

by MPG Books Ltd, Bodmin, Cornwall

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on

Blackwell Publishing, visit our website:

www.blackwellpublishing.com

IN MEMORY OF RICHARD MONTAGUE

Contents

Preface	ix
Acknowledgements	x
1. AN ANALYSIS OF COUNTERFACTUALS	
1.1 Introduction	1
1.2 Strict Conditionals	4
1.3 Variably Strict Conditionals	13
1.4 The Limit Assumption	19
1.5 'Might' Counterfactuals and Outer Modalities	21
1.6 Impossible Antecedents	24
1.7 True Antecedents	26
1.8 Counterfactual Fallacies	31
1.9 Potentialities	36
2. REFORMULATIONS	
2.1 Multiple Modalities	44
2.2 Propositional Quantification	45
2.3 Comparative Similarity	48
2.4 Similarity Measures	50
2.5 Comparative Possibility	52
2.6 Cotenability	57
2.7 Selection Functions	57
2.8 The Selection Operator	61
3. COMPARISONS	
3.1 The Metalinguistic Theory: Implicit Premises	65
3.2 The Metalinguistic Theory: Factual Premises	68
3.3 The Metalinguistic Theory: Laws of Nature	72
3.4 Stalnaker's Theory	77

4. FOUNDATIONS

- 4.1 Possible Worlds 84
- 4.2 Similarity 91

5. ANALOGIES

- 5.1 Conditional Obligation 96
- 5.2 'When Next' and 'When Last' 104
- 5.3 Contextually Definite Descriptions 111

6. LOGICS

- 6.1 Completeness Results 118
- 6.2 Decidability Results 134
- 6.3 Derived Modal Logics 137

Appendix: Related writings by David Lewis 143

Index 149

Preface

The principal changes in this revised printing are in Section 6.1, where I have corrected two major errors in my discussion of completeness results for the V -logics. Both of them were spotted by Erik C. W. Krabbe in 1976. I am most grateful to him for finding the trouble, and also for very helpful correspondence about alternative methods of repair. One error was in my construction of the canonical basis on pages 127–130: I falsely claimed that the set of co-spheres of cuts around a given index would be closed under unions.* In order to ensure such closure, it is necessary to construct the canonical basis differently. The other was in the axiom system for VC given on page 132. I left out the rule of Interchange of Logical Equivalents; however I tacitly appealed to this rule in proving completeness, so my proof did not apply to the axiom system I had given.

In addition I have corrected minor errors on pages 35, 55 and 129, also spotted by Krabbe; removed misprints; and brought some references up to date.

I have had more to say about counterfactuals and related matters. These further thoughts might appropriately have been added to this book; but since they are to be found elsewhere, I have been content to add an appendix giving citations and abstracts.

David Lewis
1986

* Erik C. W. Krabbe, 'Note on a Completeness Theorem in the Theory of Counterfactuals', *Journal of Philosophical Logic* 7 (1978): 91–93.

Acknowledgements

I am grateful to Kit Fine, Hans Kamp, David Kaplan, Richard Montague, J. Howard Sobel, Robert Stalnaker, Richmond Thomason, and many other friends and colleagues for encouragement and for valuable discussions about counterfactuals over the last five years.

I am grateful also to the American Council of Learned Societies for financial assistance, and to Saint Catherine's College, Oxford, for hospitality, during the year when most of this book was written.

David Lewis
Princeton, June 1972

1. An Analysis of Counterfactuals

1.1 Introduction

'*If kangaroos had no tails, they would topple over*' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over. I shall give a general analysis of counterfactual conditionals along these lines.

My methods are those of much recent work in possible-world semantics for intensional logic.* I shall introduce a pair of counterfactual conditional operators intended to correspond to the various counterfactual conditional constructions of ordinary language; and I shall interpret these operators by saying how the truth value at a given possible world of a counterfactual conditional is to depend on the truth values at various possible worlds of its antecedent and consequent.

Counterfactuals are notoriously vague. That does not mean that we cannot give a clear account of their truth conditions. It does mean that such an account must either be stated in vague terms—which does *not* mean ill-understood terms—or be made relative to some parameter that is fixed only within rough limits on any given occasion of language use. It is to be hoped that this imperfectly fixed parameter is a familiar one that we would be stuck with whether or not we used it in the analysis of counterfactuals; and so it will be. It will be a relation of comparative similarity.

Let us employ a language containing these two counterfactual conditional operators:

□→

* See, for instance, Saul Kripke, 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica* 16 (1963): 83–94; Richard Montague, 'Pragmatics', in R. Klibansky, *Contemporary Philosophy* (La Nuova Italia Editrice: Firenze, 1968): 102–122, reprinted in Montague, *Formal Philosophy* (Yale University Press: New Haven, 1974); Dana Scott, 'Advice on Modal Logic', in K. Lambert, *Philosophical Problems in Logic* (D. Reidel: Dordrecht, 1970); and David Lewis, 'General Semantics', *Synthese* 22 (1970): 18–67.

read as ‘If it were the case that —, then it would be the case that . . .’,
and

$$\diamond \rightarrow$$

read as ‘If it were the case that —, then it might be the case that . . .’.
For instance, the two sentences below would be symbolized as shown.

If Otto behaved himself, he would be ignored.

Otto behaves himself $\square \rightarrow$ Otto is ignored

If Otto were ignored, he might behave himself.

Otto is ignored $\diamond \rightarrow$ Otto behaves himself

There is to be no prohibition against embedding counterfactual conditionals within other counterfactual conditionals. A sentence of such a form as this.

$$((\psi \square \rightarrow ((\chi \square \rightarrow \psi) \diamond \rightarrow \phi)) \diamond \rightarrow \chi) \\ \square \rightarrow (\phi \square \rightarrow (\psi \diamond \rightarrow ((\chi \square \rightarrow \phi) \diamond \rightarrow (\phi \square \rightarrow \psi))))$$

will be perfectly well formed and will be assigned truth conditions, although doubtless it would be such a confusing sentence that we never would have occasion to utter it.

The two counterfactual operators are to be interdefinable as follows.

$$\phi \diamond \rightarrow \psi = \text{df } \sim (\phi \square \rightarrow \sim \psi),$$

$$\phi \square \rightarrow \psi = \text{df } \sim (\phi \diamond \rightarrow \sim \psi).$$

Thus we can take either one as primitive. Its interpretation determines the interpretation of the other. I shall take the ‘would’ counterfactual $\square \rightarrow$ as primitive.

Other operators can be introduced into our language by definition in terms of the counterfactual operators, and it will prove useful to do so. Certain modal operators will be thus introduced in Sections 1.5 and 1.7; modified versions of the counterfactual in Section 1.6; and ‘comparative possibility’ operators in Section 2.5.

My official English readings of my counterfactual operators must be taken with a good deal of caution. First, I do not intend that they should interfere, as the counterfactual constructions of English sometimes do, with the tenses of the antecedent and consequent. My official reading of the sentence

We were finished packing Monday night $\square \rightarrow$ we departed Tuesday morning

comes out as a sentence obscure in meaning and of doubtful grammaticality:

If it were the case that we were finished packing Monday night, then it would be the case that we departed Tuesday morning.

In the correct reading, the subjunctive 'were' of the counterfactual construction and the temporal 'were' of the antecedent are transformationally combined into a past subjunctive:

If we had been finished packing Monday night, then we would have departed Tuesday morning.

Second, the 'If it were the case that ____' of my official reading of $\square \rightarrow$ is not meant to imply that it is not the case that _____. Counterfactuals with true antecedents—counterfactuals that are not counterfactual—are not automatically false, nor do they lack truth value. This stipulation does not seem to me at all artificial. Granted, the counterfactual constructions of English do carry some sort of presupposition that the antecedent is false. It is some sort of mistake to use them unless the speaker does take the antecedent to be false, and some sort of mishap to use them when the speaker wrongly takes the antecedent to be false. But there is no reason to suppose that every sort of presupposition failure must produce automatic falsity or a truth-value gap. Some or all sorts of presupposition, and in particular the presupposition that the antecedent of a counterfactual is false, may be mere matters of conversational implicature, without any effect on truth conditions. Though it is difficult to find out the truth conditions of counterfactuals with true antecedents, since they would be asserted only by mistake, we will see later (in Section 1.7) how this may be done.

You may justly complain, therefore, that my title 'Counterfactuals' is too narrow for my subject. I agree, but I know no better. I cannot claim to be giving a theory of conditionals in general. As Ernest Adams has observed,* the first conditional below is probably true, but the second may very well be false. (Change the example if you are not a Warrenite.)

If Oswald did not kill Kennedy, then someone else did.

If Oswald had not killed Kennedy, then someone else would have.

Therefore there really are two different sorts of conditional; not a single conditional that can appear as indicative or as counterfactual depending on the speaker's opinion about the truth of the antecedent.

* 'Subjunctive and Indicative Conditionals', *Foundations of Language* 6 (1970): 89–94.

The title ‘Subjunctive Conditionals’ would not have delineated my subject properly. For one thing, there are shortened counterfactual conditionals like ‘*No Hitler, no A-bomb*’ that have no subjunctives except in their—still all-too-hypothetical—deep structure. More important, there are subjunctive conditionals pertaining to the future, like ‘*If our ground troops entered Laos next year, there would be trouble*’ that appear to have the truth conditions of indicative conditionals, rather than of the counterfactual conditionals I shall be considering.*

1.2 *Strict Conditionals*

We shall see that the counterfactual cannot be any strict conditional. Since it turns out to be something not too different, however, let us set the stage by reviewing the interpretation of strict conditionals in the usual possible-world semantics for modality. Generally speaking, a strict conditional is a material conditional preceded by some sort of necessity operator:

$$\Box(\phi \supset \psi).$$

With every necessity operator \Box there is paired its dual possibility operator \Diamond . The two are interdefinable:

$$\Diamond\phi =^{\text{df}} \sim\Box\sim\phi, \quad \text{or} \quad \Box\phi =^{\text{df}} \sim\Diamond\sim\phi.$$

If we like, we can rewrite the strict conditional using the possibility operator:

$$\sim\Diamond(\phi \ \& \ \sim\psi).$$

Or we could introduce a primitive strict conditional arrow or hook, and define the necessity and possibility operators from that.‡

A *necessity operator*, in general, is an operator that acts like a restricted universal quantifier over possible worlds. Necessity of a certain sort is truth at all possible worlds that satisfy a certain restriction. We

* Notation: sentences of our language are mentioned by means of lower-case Greek letters ϕ, ψ, χ *et al.*; sets of sentences by means of Greek capitals. Logical symbols and the like are used autonomously, and juxtaposition of names of expressions signifies concatenation of the expressions named. Possible worlds are mentioned by means of the lower-case letters h, i, j, k ; sets of worlds by means of capital letters; and sets of sets of worlds by means of script capitals.

‡ In this section only, I use the unmarked box and diamond to stand for *any* arbitrary paired necessity operator and possibility operator. When next they appear, in Section 1.5, they will be reserved thenceforth for a specific use: they will be the ‘outer’ necessity and possibility operators definable in a certain way from the counterfactual (or they will be analogously related to operators analogous to the counterfactual). The dotted box and diamond, \Box and \Diamond , will be likewise reserved when they appear in Section 1.7.