

David Kellogg Lewis

(28 September, 1941 – 14 October, 2001)

Takashi Yagisawa

California State University, Northridge

BOOKS:

Convention: A Philosophical Study (1969; Harvard University Press), with a foreword by Willard van Orman Quine, based on his doctoral dissertation (“Conventions of Language”);

Counterfactuals (1973 [revised printing 1986]; Blackwell & Harvard University Press), excerpt (as “Possible Worlds”) in Michael J. Loux [ed.] *The Possible and the Actual* (1979; Cornell University Press);

Philosophical Papers Volume I (1983; Oxford University Press);

On the Plurality of Worlds (1986; Blackwell), based in part on the 1984 John Locke Lectures at the University of Oxford;

Philosophical Papers Volume II (1986; Oxford University Press);

Parts of Classes (1991; Blackwell), with an appendix on pairing by John P. Burgess, A.P. Hazen, & David Lewis;

Papers in Philosophical Logic (1997; Cambridge University Press);

Papers in Metaphysics and Epistemology (1999; Cambridge University Press);

Papers in Ethics and Social Philosophy (1999; Cambridge University Press).

SELECTED PERIODICAL PUBLICATIONS -- UNCOLLECTED:

“Immodest Inductive Methods” (*Philosophy of Science* 38 (1971), 54-63);

“Completeness and Decidability of Three Logics of Counterfactual Conditionals” (*Theoria* 37 (1971), 74-85);

“Spielman and Lewis on Inductive Immodesty” (*Philosophy of Science* 41 (1974), 84-85);

“Possible-World Semantics for Counterfactual Logics: A Rejoinder” (*Journal of Philosophical Logic* 6 (1977), 359-63);

“Vague Identity: Evans Misunderstood” (*Analysis* 48 (1988), 128-30);

“Counterpart Theory, Quantified Modal Logic, and Extra argument Places” (*Analysis* 53 (1993), 69-71);

“Zimmerman and the Spinning Sphere” (*Australasian Journal of Philosophy* 77 (1999), 209-12);

“Causation As Influence” (*Journal of Philosophy* 97 (2000), 182-97);

“Sleeping Beauty: Reply to Elga” (*Analysis* 61 (2001), 171-76);

“Truthmaking and Difference-Making” *Noûs* 35 (2001), 602-15);

“Forget About the ‘Correspondence Theory of Truth’” (*Analysis* 61 (2001), 275-80);

“Marshall and Parsons on ‘Intrinsic’” with Rae Langton (*Philosophy and Phenomenological Research* 63 (2001), 353-55);

“Redefining ‘Intrinsic’” (*Philosophy and Phenomenological Research* 63 (2001), 381-98).

David Kellogg Lewis was the last notable systematic philosopher of the twentieth century. He is

known for his lucid writing style and numerous important contributions, broadly materialistic and Humean, to the profession in diverse areas of philosophy including his audacious advocacy of the thesis of plurality of possible worlds. Born in Oberlin, Ohio on 28 September, 1941, Lewis went to Swarthmore College in Pennsylvania and majored in chemistry before switching to philosophy. He graduated from Swarthmore with a bachelor's degree in philosophy in 1962. He then studied under Willard van Orman Quine at Harvard University and obtained a Ph.D. from Harvard in 1967. He taught at University of California, Los Angeles from 1966 to 1970 and then at Princeton University from 1970 until his untimely death in 2001.

Lewis started to exert his philosophical influence on the profession early. He wrote "An Argument for the Identity Theory" (*Journal of Philosophy* 63 (1966), 17-25, reprinted in his *Philosophical Papers Volume I*) when he was a graduate student. It was one of his three earliest publications and is now considered a classic in philosophy of mind. In the early 1960s the strongest argument for the mind-body identity theory was that given what physical sciences told us, identifying the mental with the physical gave the most economical and reasonable account of the mental. Lewis set out to provide an argument which was more direct and stronger than such an inference to the best explanation.

The first premise of his argument is one of the earliest and most concise formulations of the position known as (causal) functionalism in philosophy of mind: Any mental occurrence (Lewis uses the word 'experience') is conceptually definable in terms of its characteristic causal functional role *R* that includes typical behavioral manifestations as well as typical interactions with other mental occurrences. Such a functionalist definition is intended to tell us what it is for any particular mental occurrence to fall under a given mental type (say, perceiving that it is raining).

R is a causal role type and is specified in purely causal terms, i.e., in terms of causal relations among various internal states of the subject and also in terms of causal relations between the subject's internal states and the subject's outward physical behavior. Any particular mental occurrence falls under a given mental type by occupying *R*. Different mental occurrence types are defined by different *Rs*. Each *R* is defined by means of its relations to the other *Rs* as well as its typical perceptual input and behavioral output. This means that different mental occurrence types are inter-defined all at once. For example, one's belief that it is raining is that internal state which is typically caused by perceiving falling rain drops and which typically causes one to open an umbrella over one's head given one's desire not to get wet. At the same time, one's desire not to get wet is that internal state which typically causes one to open an umbrella over one's head given one's belief that it is raining. Mentality as a whole is thus reduced to a matter of rather complex causal organization and each type of mental occurrence is defined as a nexus in this web of causal functional whole. One implication of this functionalist picture is that it is impossible to possess just one type of mental state, say belief, without possessing the other types of mental states, say desire. Mentality is a wholesale business. Another implication of functionalism is that mentality, as such, does not dictate the material substance of which a subject of mentality has to be constituted. This opens up room for subjects of mentality that are made of a wide variety of material, including silicon chips. Thus, functionalism provides theoretical foundations for the conceptual possibility of artificial intelligence.

The second premise of Lewis' materialist argument says that according to physical sciences, which provide a true and exhaustive account of all physical phenomena, what occupies *R* is a neural phenomenon. It then follows that any particular mental occurrence of any type is a

neural phenomenon. Lewis thus argues from the type-type functionalist identity thesis, which identifies each mental type with some functional type, to the token-token materialist identity thesis, which identifies each particular occurrence of each mental type with some particular occurrence of some physical type. If Lewis is right, a suitable neurophysiological theory *implies* the identity theory instead of merely making it a reasonable hypothesis. This strengthens the materialist position significantly.

Lewis' functionalism is known as analytic functionalism. It should be distinguished from functionalism regarded as an empirical and contingent theory. Analytic functionalism is a strong philosophical thesis which asserts conceptually necessary connections between mental concepts and functional concepts, whereas the other form of functionalism is merely a natural scientific hypothesis. Thus Lewis' dedication to conceptual analysis as the backbone of philosophy was manifest already at the earliest stages of his philosophizing. He elaborated on details of his analytic functionalist position in the sequel, "Psychophysical and Theoretical Identifications" (*Australasian Journal of Philosophy* 50 (1972), 249-58, reprinted in *Papers in Metaphysics and Epistemology*, 248-61), and many other articles, most of which are included in *Philosophical Papers Volume I* as well as in *Papers in Metaphysics and Epistemology*.

Lewis reworked his doctoral dissertation into his first book, *Convention: A Philosophical Study*, which was published from Harvard University Press in 1969, when he was 28. It won the Franklin J. Matchette Prize in Philosophy, a prestigious award for scholars under 40 years old. The root motivation for the book was a defense of analyticity. The distinction between analytic and synthetic statements was made originally by Immanuel Kant and was sharpened by the so-called analytic philosophers of the twentieth century. Rudolph Carnap drew the attention of the

philosophers to a close link between analyticity and conventionality of language. Then Quine came. He launched an orchestrated attack on both analyticity and conventionality of language in his influential piece “Truth by Convention” (1936), his most famous article “Two Dogmas of Empiricism” (1951), his widely-read book *Word and Object* (1960), and “Carnap and Logical Truth” (1963). Morton White joined the attack on analyticity in his “The Analytic and the Synthetic: An Untenable Dualism” (1950). The Quine-White assault was proving rather influential, persuading many philosophers into the belief that the analytic/synthetic distinction was untenable, and in particular, that the conventions of language were a myth. Lewis’ aim in *Convention* was to resurrect a Carnapian defense of analyticity as truth guaranteed by the conventions of language. To do so, Lewis first set out to provide a general definition of convention, which went far beyond anything Carnap had produced, and then applied it to language so as to make sense of the conventions of language in particular. The book contains carefully laid out extensive lines of reasoning culminating in his final proposal, with numerous informative examples along the way to help the reader.

Lewis’ basic idea, without further refinement he puts it through to arrive at his final formulation, is that a regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a *convention* if and only if, in any instance of S among members of P , (1) everyone conforms to R , (2) everyone expects everyone else to conform to R , and (3) everyone prefers to conform to R on condition that the others do, since S is a coordination problem and uniform conformity to R is a coordination equilibrium in S . Lewis borrowed the two key concepts, coordination problem and coordination equilibrium, from game theory. When two or more agents act and the outcomes jointly depend on the actions of all the agents while the

agents' interests more or less coincide, we have a coordination problem. A coordination equilibrium is a combination of the agents' actions in which no one would be better off if any one agent acted otherwise. The main strength of Lewis' analysis is that it allows a convention to exist without any explicit agreement by the participants. It is therefore well suited to apply to language. The language of a population is a conventional signaling system for that population under a convention of truthfulness (telling the truth in that signaling system) sustained by a common interest in communication.

What of analyticity? Lewis gives a Carnapian definition for an abstract language specified in formal logical terms: A sentence is *analytic* in a language *L* if and only if the truth condition for it in *L* holds in every possible world. (Lewis would go on to use possible worlds in a number of different philosophical projects fruitfully and offer his own non-Carnapian realist metaphysics of possible worlds. More on this shortly.) So if a sentence is analytic in *L* and *L* is the actual language of our own population, then we say, the sentence is analytic. Quine, who wrote the foreword for the book, remained unconvinced by this defense of analyticity, but many others were. Thus Lewis may safely be said to have at least stemmed the surging tide of the Quine-White assault against analyticity. Lewis wrote many articles in which he further elaborated on various aspects of the book but two of them stand out: "General Semantics" (1970), "Languages and Language" (1975), and "Scorekeeping in a Language Game" (1979), all of which are included in *Philosophical Papers Volume I*.

Lewis made a number of significant contributions in the area of semantics for natural language. His "General Semantics" is a classic in the tradition of possible-worlds semantics couched in a categorial grammar. It provides a general theoretical framework for a wide variety

of grammatical constructions in a logically rigorous way. But his even better-known work in this area is his second book *Counterfactuals*, published from Harvard University Press in 1973. It contains a thorough treatment of counterfactual conditionals, various reformulations of the proposed treatment, comparisons with other theories, a metaphysical sketch of possible worlds, examinations of analogous linguistic constructions, and proofs of important formal logical results. Counterfactual (or subjunctive) conditionals are statements of the form “If it were the case that P , then it would be the case that Q ”. They are crucially important in diverse areas of philosophy. The best example of its significance is the notion of causation. David Hume made it clear that counterfactual conditionals are integral to our notion of causation. To say that C caused E is to say, at least in part, that if C had not occurred, E would not have occurred. If we do not understand counterfactual conditionals, we do not understand causation. Another example is disposition. Fragility and solubility are dispositional states. For a thing to be fragile is for it to be such that if it were impacted with force (of a small yet sufficient magnitude), it would break. For a thing to be soluble in water is for it to be such that if it were placed in water (that is not saturated), it would dissolve. Some philosophers have even argued that free will should be understood as a dispositional state: For one to be free in doing X is for one to do X and be such that if one chose to do otherwise, one would do otherwise.

Conditions under which counterfactual conditionals are true were not well understood until Robert Stalnaker’s “A Theory of Conditionals” (in N. Rescher ed., *Studies in Logical Theory*, Blackwell, 1968). Stalnaker put the theoretical apparatus of possible worlds to an elegant semantic use. According to him, ‘*If kangaroos had no tails, they would topple over*’ is true if kangaroos topple over at the closest possible world at which kangaroos have no tails.

Lewis starts the book with his typical straightforward clarity, summarizing the basic idea behind any Stalnaker-type analysis:

‘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.

Lewis’ theory is no mere rehash of Stalnaker’s theory, however. Lewis provides a much more general and thorough discussion of counterfactual conditionals than Stalnaker, subsuming Stalnaker’s analysis as a special case of his own analysis. Consider various possible worlds. Some of them are more similar to our actual world than others are. Let us imagine possible worlds as scattered around in a three-dimensional space defined by three mutually orthogonal axes X, Y, and Z, with our actual world occupying the origin. Let us also suppose that all possible worlds in this space are arranged according to the degree of similarity to our actual world. The more similar a world is to our actual world, the closer it is to the origin. The actual world is the most similar to itself, and its distance to the origin is zero. Now imagine that there is a sphere of similarity around our actual world such that at every possible world within the sphere, whenever kangaroos lack tails, they topple over—that is, there is no possible world within the sphere at which kangaroos lack tails but do not topple over. In other words, some world at which kangaroos lack tails and topple over is more similar to our actual world than any world at which kangaroos lack tails but do not topple over. Then, we would say, the counterfactual conditional *‘If kangaroos had no tails, they would topple over’* is true. This is the core of Lewis’ analysis.

Unlike Stalnaker, he does not assume that there must be a unique closest world, or even that there must be at least one closest world.

The virtues of Lewis' analysis are numerous. An obvious virtue it shares with Stalnaker's analysis is that the vagueness of counterfactual conditionals is explained by the vagueness of the similarity relation. Lewis' analysis also explains why the fallacy of strengthening the antecedent is indeed fallacious. From the statement that if Otto had come to the party, it would have been lively, it does not follow that if Otto and Anna had come to the party, it would have been lively. This is explained by means of a shifting similarity sphere. The addition '*Anna had come*' has the effect of disqualifying the previously close worlds as no longer close, opening up room for worlds at which the party was not lively. Another famous fallacy that can be explained cleanly by Lewis' analysis is the fallacy of transitivity. From the statement that if J. Edgar Hoover had been a Communist, he would have been a traitor, and the statement that if J. Edgar Hoover had been born in Russia, then he would have been a Communist, it does not follow that if J. Edgar Hoover had been born in Russia, he would have been a traitor. The similarity relations governing the two premises need not carry over to a similarity relation that makes the conclusion come out true. Another virtue of Lewis' theory is that in conjunction with the counterpart theory he proposed in "Counterpart Theory and Quantified Modal Logic" (*Journal of Philosophy* 65, 1968, 113-26, and reprinted in *Philosophical Papers Volume I*, 26-39), it can explain the simultaneous truth of Nelson Goodman's sentences, '*If New York City were in Georgia, New York City would be in the South*' and '*If Georgia included New York City, Georgia would not be entirely in the South*'. Various counterparts of New York City and Georgia in various worlds are involved, and a less stringent counterpart relation is summoned up by the subject terms '*New York City*' in (1) and

'Georgia' in (2) than by the object terms 'Georgia' in (1) and 'New York City' in (2). The result is that these two sentence are allowed to be true at the same time without difficulty. For an even more dramatic illustration of the strength of Lewis' apparatus,

take '*If I were you, ...*'. The antecedent-worlds are worlds where you and I are vicariously identical; that is, we share a common counterpart. But we want him to be in *your* predicament with *my* ideas, not the other way around. He should be your counterpart under a counterpart relation that stresses similarity of predicament; mine under a different counterpart relation that stresses similarity of ideas. (*Counterfactuals*, 43)

(I shall explain Lewis' counterpart theory more closely below.)

The most significant follow-up article to *Counterfactuals* is "Counterfactual Dependence and Time's Arrow" (*Noûs* 13, 1979, 455-76, reprinted in *Philosophical Papers Volume II*, 32-52), in which he responds to Kit Fine's objection in "Critical Notice of Counterfactuals by D. Lewis" (*Mind* 84, 1975, 451-58). Fine argues that since a world without a nuclear holocaust is closer to our actual world than a world with a nuclear holocaust, the counterfactual conditional '*If Nixon had pressed the button, there would have been a nuclear holocaust*' comes out false on Lewis' analysis. The notion of similarity was left vague and contextually fluid in the book. Fine's objection exploits that. Lewis therefore responds by giving more detailed standards of similarity between worlds that are appropriate for assessing counterfactual conditionals. In the order of decreasing importance, they are: (1) avoidance of big, widespread, diverse violations of law, (2) maximization of the spatio-temporal region in which a perfect match of particular fact prevails, (3) avoidance of small, localized simple violations of law, and (4) securing approximate similarity of particular fact. According to these standards of similarity, a world where Nixon pressed the

button and a nuclear holocaust occurred is more similar to our actual world than any world where Nixon pressed the button and a nuclear holocaust did not occur.

Among the many philosophical theses Lewis articulated and defended throughout his short yet prolific life, the most celebrated, or infamous, one of all is his thesis of the plurality of worlds, or modal realism. It is the thesis that our world is but one among many worlds each of which is as real and concrete as our world. He argues for it in a number of places but the most important are “Anselm and Actuality” (*Nous* 4, 1970, 175-88, reprinted in *Philosophical Papers Volume I*, 10-21), “Counterpart Theory and Quantified Modal Logic”, *Counterfactuals* (especially 4.1 Possible Worlds), and *On the Plurality of Worlds*, which was based partly on the prestigious John Locke Lectures he delivered at Oxford University in 1984 and was published by Blackwell in 1986. By ‘our world’ Lewis means the totality of absolutely everything that bears any spatiotemporal relation to us here and now. Modal realism says, in addition to our world, there are many other worlds and each of these worlds is of the same kind of entity as our world. Each of them is an inclusive totality of spatiotemporally related things. Since anything that bears any spatiotemporal relation to a thing in a world is also a thing in that world, it follows that no two worlds stand in any spatiotemporal relation to each other. Every world is spatiotemporally completely isolated from every other world. This means, among other things, that there are absolutely no causal interactions between any two worlds. This in turn means that traveling from one world to another is impossible. However far one traveled by whatever fancy means, one would never get out of one’s own world.

How many such spatiotemporally and causally isolated worlds are there altogether? Based on a bit of technical reasoning, Lewis estimates that there are more worlds than there are real

numbers. An obvious objection to such an extravagant ontology is readily suggested by the principle of parsimony known as Ockham's Razor: "Do not multiply entities beyond necessity". Lewis replies by distinguishing two versions of Ockham's Razor and saying that one version is not worth observing and that his theory observes the other, worthy version. The unworthy version says that we should keep the total number of *individual* entities postulated in our theory low. Lewis sees no plausibility in it. The worthy version says that we should keep the total number of *kinds* of entity postulated in our theory low. Since the other worlds are of the same kind of entity as our world, Lewis does not violate the principle by postulating them, assuming that it is warranted to believe in the existence of our own world. Lewis also believes that postulating the other worlds is not "beyond necessity". He believes that his thesis of modal realism provides us with the best overall philosophical outlook. Lewis puts it this way:

One comes to philosophy already endowed with a stock of opinions. It is not the business of philosophy either to undermine or to justify these preexisting opinions, to any great extent, but only to try to discover ways of expanding them into an orderly system. A metaphysician's analysis of mind is an attempt at systematizing our opinions about mind. It succeeds to the extent that (1) it is systematic, and (2) it respects those of our pre-philosophical opinions to which we are firmly attached. Insofar as it does both better than any alternative we have thought of, we give it credence. There is some give-and-take, but not too much: some of us sometimes change our minds on some points of common opinion, if they conflict irremediably with a doctrine that commands our belief by its systematic beauty and its agreement with more important common opinions.

So it is throughout metaphysics; and so it is with my doctrine of realism about

possible worlds. Among my common opinions that philosophy must respect (if it is to deserve credence) are not only my naïve belief in tables and chairs, but also my naïve belief that these tables and chairs might have been otherwise arranged. Realism about possible worlds is an attempt, the only successful attempt I know of, to systematize these preexisting modal opinions. (*Counterfactuals*, 88)

This passage gives a reason for embracing modal realism but it goes far beyond that. It succinctly summarizes Lewis' view on the nature of philosophy as an adjudicating and systematizing endeavor.

The example of tables and chairs illustrates a core application of the thesis of modal realism within the topic of alethic modality, which deals with various modalities of truth, such as actual truth, possible truth, necessary truth, contingent truth, impossible truth, etc.: e.g., “It might have been true that the tables and chairs be positioned differently”; “It could not be true that π be a rational number”. Alethic modality should be distinguished from other types of modality, especially, epistemic modality: e.g., “As far as I know, the tables and chairs might be positioned differently”; “For all little Davy knows, π could be rational”. Alethic modality is independent of epistemic modality, but epistemic modality is definable in terms of alethic modality: to say that such-and-such is epistemically possible is to say that what one knows and such-and-such could be jointly true, and to say that such-and-such is epistemically necessary is to say that what one knows and the negation of such-and-such could not be jointly true. Other types of modality are also amenable to definitions in terms of alethic modality. Thus, alethic modality is the most basic type of modality.

We believe that tables and chairs exist in certain spatial relationships to one another and to

other things in our world. We also believe that they might possibly have been arranged in any number of different spatial configurations. But what does it mean to say of the tables and chairs that they might possibly have been arranged in a spatial configuration in which they are not in fact arranged? We believe that Al Gore existed in 2000 and lost the U. S. presidential election then in our world. We also believe that he might possibly have won the election. What does it mean to say of Al Gore that he might possibly have won the 2000 election which he did not in fact win? For anyone who takes talk of possible worlds with any philosophical seriousness, to say that the tables and chairs might possibly have been arranged in a certain non-actual configuration is to say that there is some possible but non-actual world according to which the tables and chairs are arranged in that spatial configuration, and to say that Al Gore might possibly have won the 2000 election is to say that there is some possible but non-actual world according to which Al Gore won the election. Lewis agrees with this, of course. But how does he understand the locution, “according to a world, such-and-such is the case”? Here he speaks of *counterparts*. To say that according to a world W_1 , the tables and chairs are arranged in a non-actual way is to say that W_1 contains counterparts of the tables and chairs and those counterparts are arranged in that non-actual way in W_1 . To say that according to a world W_2 , Al Gore won the 2000 election is to say that W_2 contains a counterpart of Al Gore and that counterpart won the election in W_2 . For Lewis, not only is every world spatiotemporally isolated from every other world, but also no concrete object existing in any one world exists in any other world. That is, every possible concrete object exists in one world and one world only. Thus, Lewis analyzes the possibility of the tables and chairs’ being arranged in a non-actual configuration in terms of the existence of a world in which some other tables and chairs, counterparts of the original tables and chairs, exist

and are arranged in that configuration. Likewise, he analyzes the possibility of Al Gore's winning the election in terms of the existence of a world in which some other man, a counterpart of Al Gore, exists and wins the election. What makes one possible thing in one world a counterpart of another possible thing from another world? Lewis understands the counterpart relation as a kind of similarity relation.

This part of his doctrine, known as the counterpart theory, is one of the most controversial aspects of his metaphysics and has generated an enormous amount of literature. Some of the objections made against Lewis' counterpart theory are simply based on misunderstanding of the theory. One commonly made objection, for example, says that a victory by someone other than Al Gore in some world is irrelevant to the possibility of Al Gore's victory. After all, the objector would say, it is Al Gore himself that we are talking about when we say that Al Gore might possibly have won. Some other man, looking very much like Al Gore, doing similar things as Al Gore, and even bearing the name 'Al Gore', may be the winner of the election in some other world, but what does that have to do with the possibility of a victory by Al Gore, that very man? This objection is confused. To see that, remember that anyone who takes the theoretical apparatus of possible worlds seriously as useful for analyzing alethic modality would say that the possibility of Al Gore's victory is to be analyzed in terms of the existence of a possible world according to which Al Gore wins. Lewis emphatically agrees with this. For Lewis, just as for any serious possible-worlds theorist, Al Gore might have won if and only if according to some world Al Gore did win. It is Al Gore himself who is said to have won according to some world. Lewis goes on to elaborate what we mean when we say, "Al Gore won according to some world"; to say that Al Gore won according to some world is to say that Al Gore has a counterpart in that

world and that counterpart won in that world. Al Gore's counterpart in such a world is not Al Gore but is still a counterpart of *Al Gore*. So, it is Al Gore, that very man, who is said to have won according to a non-actual world.

In *On the Plurality of Worlds*, he collects and discusses other theoretical gains to be had by modal realism in addition to the analysis of alethic modality, counterfactual conditionals, and linguistic meanings. The definition of a property as a set of all of its possessors is a case in point. The traditional nominalist position on the nature of properties is that they are nothing above and beyond particular objects that possess them. If we combine this with the claim that everything is actual, we have a problem. We will not be able to distinguish two different yet accidentally co-extensional properties. That is, we will be forced to say that whenever a property P_1 and a property P_2 are possessed by exactly the same particular objects, then P_1 and P_2 are one and the same property. This is unsatisfactory, for even if all purple monsters are man-eating monsters and all man-eating monsters are purple monsters, it should not follow that the property of being a purple monster and the property of being a man-eating monster are one and the same property. Lewis proposes to uphold the traditional nominalist position by separating it from the claim that everything is actual. If we identify a property with the set of its possessors and claim at the same time that the possessors may be actual or merely possible objects, we can escape the problem. Even if all and only purple monsters are man-eating monsters in one world, there are many other worlds where some purple monsters are not man-eating monsters and many other worlds where some man-eating monsters are not purple monsters. So, the set of purple monsters is different from the set of man-eating monsters.

It is a mistake to object to Lewis' modal realism by saying that only our world and its

inhabitants actually exist. First of all, Lewis will agree that only our world and its inhabitants actually exist. Secondly, modal realism says that there are more things than what actually exists. To substantiate these points, Lewis offers what is called the “indexical analysis of actuality”. To say of something that it actually exists, according to Lewis, is to say that it exists in our world. It is analogous to saying that something exists here. The word ‘here’ is a stereotypical example of so-called “indexical” words. Its temporal analog is the word ‘present’. ‘Here’ means the spatial location of the speaker, and ‘present’ means the temporal location of the speaker. When Stephanie says, “A donkey is here”, what she is saying is that a donkey is around where she is. No matter where she is, if there is a donkey around where she is, what she says is true. When Stephanie says, “The starting time of the test is present”, what she is saying is that the test starts as she speaks. No matter when she speaks, if the test starts at the time of her speaking, what she says is true. Likewise, according to the indexical analysis of actuality, when Stephanie says, “A talking donkey actually exists”, what she is saying is that a talking donkey exists in the world in which she speaks. No matter in which world she speaks, if a talking donkey inhabits the world in which she speaks, what she says is true. Of course, if she speaks in our world, what she says is (very likely) false. There is a sense in which every place is “here” for anyone who occupies it and a sense in which every time is “present” for anyone who persists through it. Analogously, there is a sense in which every world is “actual” for anyone who inhabits it. In this sense, ‘actual’ means ‘this-worldly’. Lewis’ indexical analysis of actuality offers us a straightforward answer to the age-old conundrum: “Why am I actual rather than merely possible?” The answer is: “Because that is what ‘actual’ means.” Whenever I call myself “actual”, I am bound to be right, just as whenever I say “I am here”, I am bound to be right. I inhabit a world I inhabit. “I am actual” is

as analytic and knowable *a priori* as “I am here”.

Some might be tempted to identify Lewis’ modal realism with the interpretation of quantum physics known as the “many-worlds” interpretation, which postulates that reality splits into many different alternative “worlds” every time a measurement is made. It is a mistake to identify Lewis’ theory with this interpretation of quantum mechanics. Lewis did not propose his ontology as an interpretation of quantum mechanics, or as an interpretation of any particular scientific theory for that matter. His ontology is an integral part of a wide-ranging metaphysical theory that provides conceptual analyses of alethic modality and other philosophically important notions. Lewis’ possible worlds are fundamentally different from the “worlds” of the “many-worlds” interpretation of quantum mechanics. The latter bear basic spatiotemporal relations to one another; different “worlds” split away from a common “world”. Lewis’ possible worlds, by contrast, bear no spatiotemporal relation of any kind to one another. If something bears any spatiotemporal relation to a given possible world, then that thing is a part of that possible world, and it is not another possible world. No two Lewisian possible worlds overlap in any way, hence no two Lewisian possible worlds share an initial segment and then split away. The “many-worlds” interpretation of quantum mechanics is often believed to have been proposed by Hugh Everett, III. But this is a mistake. Everett’s interpretation of quantum mechanics does not postulate splitting “worlds” but instead takes superposition as actual and prevalent. This complicates our view of what the actual world is like but does not require other “worlds”. Thus, Lewis’ theory has nothing to do with either the “many-worlds” interpretation of quantum mechanics or Everett’s interpretation of quantum mechanics.

Many modal metaphysicians agree with Lewis on the philosophical usefulness of speaking

of possible worlds. They agree that alethic modality and counterfactual conditionals are well analyzed in terms of possible worlds. However, they strongly disagree with Lewis on the nature of possible worlds. Unlike Lewis, who regards worlds as concrete spatiotemporal wholes, they regard possible worlds as abstract representations of some kind or other. Lewis calls such theories *ersatz modal realism*. Ersatz modal realists help themselves to the conceptual apparatus of possible worlds in their endeavor to clarify and advance philosophical debates, but they steadfastly refuse to embrace the Lewisian ontology. For them, possible worlds and their inhabitants are useful tools in philosophizing about difficult issues and attaining conceptual and logical clarity but do not carry serious ontological weight. Ersatz modal realists wish to have the cake of possible worlds and eat it, too. They refuse to postulate concrete possible worlds as Lewis does but instead propose shadowy replacements for them. That is why Lewis calls them “ersatz” modal realists. Lewis has two main objections against ersatz modal realism. The first objection is that ersatz modal realism assumes the notion of possibility rather than providing an analysis of it. Take a linguistic version of ersatz modal realism, for example, according to which each Lewisian possible world is replaced by a maximal consistent set of sentences. This yields the analysis of possibility which says that it is possible that P if and only if there is a maximal consistent set of sentences according to which it is true that P . Lewis’ objection in a nutshell is that this is circular. Possibility is analyzed in terms of the consistency of a set of sentences. But the consistency of a set of sentences is best understood as a possibility of the truth of all sentences in that set. So, possibility is analyzed ultimately in terms of possibility, which is viciously circular. Most sympathizers of ersatz modal realism find this objection devastating enough to retreat to the position that use of ersatz worlds should be regarded as nothing more than a useful heuristic in

philosophical theorizing. The second objection Lewis has against ersatz modal realism is that it lacks sufficient resources to represent all distinct possibilities as genuinely distinct. There are a variety of ersatz modal realism but all of them draw on the resources of reality as it actually is for constructing the abstract representations of non-actual possibilities. This unduly limits the expressive power of the resultant representations. For example, there might possibly have been more fundamental properties than there actually are, that is, more physical properties as basic as charge or spin but not reducible to actual fundamental properties. If we confine our resources to actuality, we are unable to construct any representations for such alien basic properties. Ersatz modal realists have attempted some replies but none has proven satisfactory. There is even a convincing argument that ersatz modal realism is logically inconsistent (Grim, *The Incomplete Universe*, 1991).

The fourth and last book Lewis wrote, *Parts of Classes*, is a treatise on the foundations of mathematics. Set theory provides a unified basis for almost all of mathematics, and therefore the question concerning the nature of sets lies at the heart of philosophy of mathematics. The orthodox view of sets takes them to be abstract objects with no spatiotemporal location. The problem is that the relation between a set and its members is totally obscure. Take a particular chess board. We may see it as a whole consisting of nine rows, each of which contains nine squares. Alternatively, we may see it as a whole consisting of nine columns, each of which also contains nine squares. For any collection of particular objects, there is a set that has precisely those objects as its members. So, there is a set whose members are precisely the nine rows, and there is a set whose members are precisely the nine columns. Since no row is a column (and hence no column is a row), these sets share not a single member in common. So, they are two

entirely distinct sets. At the same time, the whole consisting of the rows as its parts and the whole consisting of the columns as its parts are one and the same object, namely, the entire chess board. Therefore, neither set is a whole consisting of its members as parts. In general, a set is not a whole consisting of its members as parts. Lewis' central thesis in the book is that a set is a whole consisting of its subsets as parts. Since every non-empty set can be partitioned into singletons (sets with just one member), Lewis' central thesis is tantamount to the claim that every set is a whole consisting of the singletons of its members as parts. The exception is the empty set, the set with no member. Lewis identifies the empty set as the whole consisting of all things that are not sets. By advocating his central thesis, Lewis took a significant first step toward subsuming mathematics under mereology, a general theory of the part-whole relation. What requires further investigation is the nature of the relationship between a singleton and its sole member. The book's appendix, which Lewis wrote with his Princeton colleagues, John Burgess and Gideon Rosen, addresses this vexing question.

Lewis was a train buff. He enjoyed traveling by train. He would ride for hours, while reading and writing philosophy. He had an extensive model train set in the basement of his house in Princeton. He was also very fond of Australia and Australians. He would routinely spend his summer break there in Australia's winter. The philosophical influence on Lewis by Australian philosophers, especially David M. Armstrong, was evident from the beginning of his academic career to the very end. At the same time, the philosophical influence Lewis exerted on the Australian philosophers was immeasurable. The state of philosophy in Australia now would be drastically different and definitely inferior if it had not been for Lewis' influence. In fact, the state of analytic philosophy in the world now would be very different and certainly inferior if it had not

been for Lewis' prolific contributions of the highest quality throughout his relatively short life.

The philosophical world would have been significantly richer if he had lived even for just ten more years.

REFERENCES:

Everett, Hugh III. "'Relative State' Formulation of Quantum Mechanics", *Reviews of Modern Physics* 29 (1957), 454-62;

Fine, Kit. "Critical Notice of Counterfactuals by D. Lewis", *Mind* 84 (1975), 451-58;

Forrest, Peter and Armstrong, D. M. "An Argument Against David Lewis' Theory of Possible Worlds", *Australasian Journal of Philosophy* 62 (1984), 164-68;

Grim, Patrick. *The Incomplete Universe* (1991: MIT Press);

Hazen, Allen. "Counterpart-Theoretic Semantics for Modal Logic", *Journal of Philosophy* 76 (1979), 319-38;

Lycan, William. "The Trouble with Possible Worlds", in Loux, Michael (ed.) *The Possible and the Actual: Readings in the Metaphysics of Modality* (1979: Cornell University Press), 274-316;

Skyrms, Brian. "Possible Worlds, Physics and Metaphysics" *Philosophical Studies* 30 (1976), 323-32;

Van Inwagen, Peter. "Two Concepts of Possible Worlds", in French, P., Uehling, T., and Wettstein, H. (eds.) *Midwest Studies in Philosophy, Vol. XI: Studies in Essentialism* (1986: University of Minnesota Press), 185-213.