Humean Supervenience: Lewis (1986, Introduction)  
7 October 2010: J. Butterfield

1: Humean supervenience and the plan of battle:—
Three key ideas of Lewis’ mature metaphysical system are his notions of possible world, of natural property, and ‘Humean supervenience’. The idea of Humean supervenience is that all truths supervene on truths about matters of local particular fact: where ‘matters of local particular fact’ is to be understood in terms of Lewis’ metaphysics of natural properties, with the properties having spacetime points, or perhaps point-sized bits of matter, as instances. Thus he writes that Humean supervenience

\[\ldots\text{is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing then another. (But it is no part of the thesis that these local matters are mental.)}\]

We have geometry: a system of external relations of spatiotemporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at these points, we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without a difference in qualities. All else supervenes on that. (1986, pp. ix-x).

and similarly

\[\ldots\text{says that in a world like ours, the fundamental relations are exactly the spatiotemporal relations: distance relations, both spacelike and timelike, and perhaps also occupancy relations between point-sized things and spacetime points. And it says that in a world like ours, the fundamental properties are local qualities: perfectly natural intrinsic properties of points, or of point-sized occupants of points. Therefore it says that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past and present and future. (1994, pp. 225-226.)}\]

Humean supervenience is not widely believed—few philosophers even sign up to all the notions deployed in its statement. But it has been a natural focus of metaphysicians’ attention. In the literature, we can distinguish three broad groups of topics:

(i): Lewis’ battle-plan and its cousins: Issues about whether to analyse notions such as law of nature, causation, chance etc., and “higher-level” concepts about mind and language, in terms of the notions of Lewis’ framework. Lewis (1986, p. xi-xiv) sketches how his work on all these topics provides a “battle-plan”: i.e. roughly, a sequence of supervenience claims for these concepts.

[1]: The first step is laws of nature: which Lewis takes to be the exceptionlessly true contingent generalizations with an unexcelled combination of simplicity and strength (Counterfactuals, p. ??).
[2]: Then, counterfactuals: though true counterfactuals are made true by the qualitative nature of the actual world, Lewis claims that mention of other possible worlds, and degrees of similarity between worlds (to which laws contribute strongly) is indispensable for giving their truth-conditions. (Cf. Section 3 below.)

[3]: Then, causation: taken as a chain of counterfactual dependence between propositions stating the occurrence or non-occurrence of events.

[4]: Then, persistence through time: taken as qualitative and causal continuity between temporal parts, or stages (Cf. 21 October and (1999)).

[5]: Then, mind: analytical functionalism (details 28 October), combined with mind-brain identity theory, and e.g. assimilation of knowledge gained from the *qualia* of experience to knowledge-how rather than knowledge-that (cf. ‘What experience teaches’, reprinted in (1999a)).

[6]: Then, language: building on the ideas of Grice, Bennett and himself (in *Convention*, and e.g. ‘Radical interpretation’, cf. 11 November).

More generally, much literature of neo-Humean stripe is concerned with how truths using familiar central concepts of common-sense knowledge and belief—concepts such as law, causation, the persistence of objects, and mental and semantic concepts such as belief and reference—might supervene on a basis acceptable to Humeans, though perhaps not exactly the basis proposed by Lewis. (For example: for laws, cf. Earman and Roberts *Philosophy and Phenomenological Research* (2006).) And some of these truths pose a challenge in that they seem not to thus supervene; cf. (iii) below.

(ii): General metaphysical issues: about the notions of Lewis’ framework, in particular possible worlds and natural properties, and-or about related notions.

[a]: For example, one well-known issue is: can possible worlds and natural properties be construed less “realistically” than Lewis proposes (e.g. Taylor *Mind* 1993), and yet do the philosophical work they are meant to do?

[b]: Another example: is Lewis right in his analysis of the intrinsic-extrinsic distinction among properties in terms of his notion of perfectly natural properties? (Details on 14 October).

[c]: Another example: Humean supervenience is intended as a contingent thesis. Lewis uses his theory of natural properties to specify at which worlds he holds it true (and thus hopes the actual world is among them!): (1986, p. x; details on 14 October). But one can formulate essentially the same doctrine relative to a scientific theory, rather than relative to a possible world. (Of course a metaphysician like Lewis who accepts the idea of a law of nature can link relativizations to a theory and to a possible world using the idea of the “complete” theory of a world, say as an axiomatization of all its laws of nature.)

(iii): Direct threats to Humean supervenience. There are two main examples. First, chance; which Lewis addresses in detail in (1986, pp. xiv-xvi, 121-131), and to his greater satisfaction in (1994). Second, persistence (cf. 21 October). For Lewis as a Humean wants to be perdurantist, in his sense (1986a, Sec 4.2): this means that he faces the Leibniz-Broad-Armstrong-Kripke ‘rotating discs argument’.
2: Classifying relations as internal, or external—or neither: —
Lewis states a trichotomy among relations corresponding to the intrinsic-extrinsic dichotomy among properties: (1983a, p. 26 fn. 16; 1986a, p. 62). It is useful and has become widespread. Though he explains it in terms of his preferred understanding of intrinsic and extrinsic properties (viz. defined in terms of his natural properties), the trichotomy can be explained in the very same words, using other understandings of intrinsic and extrinsic. The trichotomy also uses the idea of the mereological fusion, or composite, of objects: an idea I am happy to accept, and for which there is a powerful argument (Lewis 1986a, pp. 212-213, developed by Sider Four dimensionalism 2001, pp. 121-139).

(1) An internal relation is determined by the intrinsic properties of its relata. So if \( xRy \), and \( x' \) matches \( x \) in all intrinsic properties, and \( y' \) matches \( y \) in all intrinsic properties, then we must have \( x'Ry' \). So any relation of similarity or difference in intrinsic respects is internal; for example, if height is an intrinsic property, then ‘being taller than’ is an internal relation.

(2) On the other hand, there are relations, notably relations of spatiotemporal distance, that are not internal, but do supervene on the intrinsic nature of the composite (mereological fusion) of the relata. Thus suppose \( x, y \) are point-particles 1 metre apart. Then it seems reasonable to say both of the following:—

(i): There could be point-particles \( x', y' \) that intrinsically match \( x \) and \( y \) respectively, and that are 2 metres apart—so that distance is not internal. But on the other hand:

(ii): Any object intrinsically matching the fusion or composite of \( x \) and \( y \) would have two parts intrinsically matching \( x \) and \( y \), 1 metre apart.

Accordingly, Lewis calls relations that supervene on the intrinsic nature of the fusion of the relata, external; and he takes (ii) to show that spatiotemporal relations are external.

(3) Finally, there are relations that do not supervene even on the intrinsic nature of the composite of the relata; i.e. relations that are neither internal nor external. Lewis’ example is the relation having the same owner: \( x \) and \( y \) could intrinsically match \( x' \) and \( y' \) respectively, and their composites might also match; and yet \( x \) and \( y \) might have the same owner, while \( x' \) and \( y' \) do not.

3: Must actual truths have actual truthmakers? What about counterfactuals?!: —
Though ‘truthmaker’ is a philosophical term-of-art awaiting strict definition, the way that Lewis’ truth-conditions for counterfactual conditionals mention possible worlds, and their relations of similarity, makes one think that—whether one takes truthmakers to be facts or objects—a counterfactual has truthmakers “scattered across the worlds”—apparently violating the surely-desirable principle that actual truths have actual truthmakers.

Lewis of course recognizes that his proposed truth-conditions for counterfactuals in terms of similarity between possible worlds threaten this principle. After all, Lewis proposes for an actually true counterfactual, truth-conditions in terms of other worlds! Thus recall that, roughly speaking, \( A \rightarrow C \) is actually true if some \((A\&C)\)-world is closer (i.e. more similar) to the actual world than any \((A\&\neg C)\)-world is. So he writes:
Here is our world, which has a certain qualitative character. (In as broad a sense of ‘qualitative’ as may be required—include irreducible causal relations, laws, chances, and whatnot if you believe in them.) There are all the various A-worlds, with their various characters. Some of them are closer to our world than others. If some (A&C)-world is closer to our world than any (A&¬C)-world is, that’s what makes the counterfactual true at our world. Now ... it’s the character of our world that makes some A-worlds be closer to it than others. So, after all, it’s the character of our world that makes the counterfactual true—in which case why bring the other worlds into the story at all?

To which I reply that it is indeed the character of our world that makes the counterfactual true. But it is only by bringing the other worlds into the story that we can say in any concise way what character it takes to make what counterfactuals true. The other worlds provide a frame of reference whereby we can characterize our world. By placing our world within this frame, we can say just as much about its character as is relevant to the truth of a counterfactual (1986a, p. 22).

This passage makes two main claims, one in each paragraph:

(Actual): although Lewis’ truth-conditions mention other worlds, it is the character of the actual world that makes the counterfactual actually true;

(Concise): mentioning other worlds is the only concise way to state what in the actual world’s character is relevant to the counterfactual’s truth.

Of these two claims, (Actual) is more important for us—it summarizes both the threat to truthmakers and Lewis’ reply.

We can better understand (Actual) by recalling Lewis’ (1986a, p. 62) distinction between (a) relations that supervene on the intrinsic properties of their relata, which Lewis calls ‘internal’, and (b) relations that do not thus supervene, which I will from now on call ‘external’. (Thus I will not need Lewis’ doctrines about which properties are intrinsic, and can make do with some intuitive if disputable examples of intrinsic properties. Nor will I need Lewis’ allowance that a relation might supervene on the composite of the relata taken together: his main example of this category being spatiotemporal relations.)

Thus relations of similarity or difference in intrinsic respects are internal; so that if an object’s mass is an intrinsic property of it, the relation ‘is more massive than’ is internal. An example of an external relation would be ‘has the same owner as’: a and a’ could match in all their intrinsic properties and yet a person might own a and some other object b, but not a’; so that ‘has the same owner as’ holds of ⟨a, b⟩ but not ⟨a’, b⟩.

Lewis applies this distinction not just to relations between objects in a single world, but to objects in different worlds. Thus a sentence such as ‘He is slimmer than he would have been without the diet’ reports an internal relation between objects in different worlds (a man and one of his counterparts). A sentence reporting a transworld external relation seems harder to construct; I suppose because our thought and language has little use for them. But Lewis’ own counterpart theory gives examples. For counterparthood, though it sometimes emphasises intrinsic similarity, often emphasises extrinsic similarity, especially as regards the object’s origins (Lewis 1986a, p.88). Thus two objects a and a’ (in the
same world, or in two different worlds) might be duplicates, while only $a$ is a counterpart of some object $b$ in another world—say an actual object $b$.\footnote{Here is an example with $a, b$ both actual, and indeed identical: ‘an atom-for-atom replica of Humphrey (as he actually was at, say, noon 4 July 1968), who had been born of different parents than the actual Humphrey (in say Latvia, never setting foot in the USA etc.), would not have been [folk-language, according to Lewis, for: would not have been a counterpart of] Humphrey’. Here, $a = b$ is the actual Humphrey, and $a'$ is the replica. Another example, with $a$ and $b$ in different worlds: ‘Each of two people might be atom-for-atom replicas of Humphrey as he actually was at noon, 4 July 1968; but only the person whose origin matched (at least: sufficiently closely) that of the actual Humphrey, would be Humphrey’. Here, $a, a'$ are the replicas, $b$ is the actual Humphrey.}

Furthermore, Lewis also takes worlds to be objects (in short: the mereological fusion of their parts) and so allows them as relata; and therefore applies this distinction to relations between worlds. And he says explicitly (1986a, p. 62, 177) that since the relation of closeness between possible worlds used in his analysis of counterfactuals is a relation of similarity, it is internal. Hence his claim in (Actual) that the truth-values of counterfactuals are determined by the character of our world. For the character of our world determines which worlds are similar to it. (Though it is a vague and controversial matter which respects of similarity are relevant to the truth-conditions of counterfactuals (‘Counterfactual dependence and time’s arrow’, Nous 1979), any resolution of those issues will render the overall similarity relation internal.)

Now it is clear how Lewis’ (Actual) agrees with the principle that actual truths have actual truthmakers. Agreed, the way that Lewis’ truth-conditions mention other worlds makes one think that—whether one takes truthmakers to be facts or objects—a counterfactual has truthmakers “scattered across the worlds”. But Lewis replies: ‘No worries: which facts, objects etc. in other worlds get mentioned in the truth-conditions is wholly determined by the character of the actual world—and that is sufficient for satisfying the idea that actual truths have actual truthmakers.’ And Lewis might well go on: ‘If you want, you can call the facts, objects etc. in the other worlds that get mentioned in the truth-conditions ‘truth-makers’. But the point remains that their being scattered across the worlds is innocuous. The fact that the character of the actual world determines them (and thereby the truth-value of the counterfactual) is sufficient to satisfy the spirit, if not the letter, of the principle that ‘actual truths have actual truthmakers’.’

References: