Problèmes de Métaphysique I et II

Séminaire MA3 2007-8

Survenance et Détermination, 10 octobre 2007

Philipp Keller

The explanandum

"Supervenience", though a philosophers' notion, has a venerable history (cf. Horgan 1993). It was used by Leibniz to say that relations are nothing over and above the intrinsic properties of their relata, by the British emergentists to characterise the special sciences (cf. McLaughlin 1992), by Sidgwick to say that moral characteristics covary with non-moral ones, by Moore (1922: 261) to say that the former are grounded in the latter, by Hare (1952: 80–81) to say that they stand in some relation of strict implication and by Davidson (1970: 214) to say that "mental characteristics are in some sense dependent, or supervenient, on physical characteristics" (cf. Kim 1990: 136–138). Here is what Robert Stalnaker (1996) says about the "intuitive ideas that motivate the attempts to articulate concepts of supervenience":

> "To say that the A-properties or facts are supervenient on the B-properties or facts is to say that the A-facts are, in a sense, redundant, since they are already implicitly specified when one has specified all the B-facts. A-facts are not fact 'over and above' the B-facts, not something 'separate'. To state an A-fact, or ascribe an A-property, is to describe the same reality in a different way, at a different level of abstraction, by carving the same world at different joints." (Stalnaker 1996: 87)

Kim (1990: 140) identifies three key features of our concept of supervenience: covariance, dependency and non-reducibility (where "non-reducibility" means that the supervenience of A-features on B-features is consistent with the former not being reducible to the latter).[1] Explanation, sometimes required for reducibility, is absent from Kim's list: supervenience claims state that some patterns of property covariation hold, without explaining why they hold:

> "The concept of supervenience is supposed to be a concept that helps to isolate the metaphysical part of a reductionist claim – to separate it from claims about the conceptual resources and explicit expressive power of theories we use to describe the world." (Stalnaker 1996: 89)

Three explanantia

A first notion of supervenience, also called the weak one, is what Jackson (1998: 9) calls *intra-world* supervenience:

Definition 1 (Intra-world supervenience)**.** *A set of properties A intra-worldly supervenes upon a set of properties B iff, for any possible world w, if x and y are B-indiscernible in w, then they are A-indiscernible in w.*[2]

[1] He now retracts the condition of non-reducibility (Kim 1993a: 165, n. 5) and holds that supervenience is a kind of reducibility after all. I find this terminological choice unfortunate: reducibility is better called "reducibility" than "supervenience". I hold with Campbell (1981: 15) that "one feature is supervenient upon others if, while not following deductively from those others, it nevertheless cannot vary unless they do", adding just the dependence component.

[2] This is the version of (Stalnaker 1996: 91) and also of Kim (1984: 64) who shows that it is equivalent to the following (assuming the base set to be closed under property negation and (possibly infinite) property conjunction):

$$\Box\forall x\forall F\in A(Fx\rightarrow\exists G\in B(Gx\wedge\forall y(Gy\rightarrow Fy))) \tag{1}$$

Intra-world supervenience (1), as (Jackson 1998: 10) notes, is clearly too weak to count as a species of determination: it does not capture relational dependencies (not even generic ones) and does therefore not secure that the B-nature of something *alone* secures its A-nature. The property of being among the tallest things intra-worldy supervenes on the individual height some object has, but something's being among the tallest things does not just depend on its height, but also on the heights of its world-mates. It falls short of the condition that "fixing the base properties of an object fixes its supervenient properties" (Kim 1984: 60) for whether or not all A-alike things have B or lack B depends on the world in question. It therefore does not support counterfactuals: we cannot say that if something had the subvening properties, it would also have the supervening ones.

Weak supervenience is clearly too weak: it does not rule out that the covariance in question is purely accidential. We therefore strengthen weak to strong supervenience and say that a set of properties A *inter-worldly supervenes* upon a set of properties B iff, for any possible individuals x and y, if they are B-indiscernible, then they are are A-indiscernible.

Instead of quantifying just over possible worlds we now quantify possible individuals, instead of indiscernibility with respect to some possible world, we now have indiscernibility *tout court*. If we say that two possible individuals are indiscernible iff they have the same properties in their respective worlds, we get:

Definition 2 (Inter-world supervenience)**.** *A set of properties A inter-worldly supervenes upon a set of properties B iff, for any worlds w and v and any individuals x and y, if x has the same B-properties in w than y has in v, then x has in w the same A-properties than y has in v.*[3]

Strong, inter-world, supervenience entails, but is not entailed by, weak, intra-world, supervenience.

The main difference between weak (1) and strong supervenience (2) is that we strengthen the notion of sameness of properties from co-extensiveness to necessary co-extensiveness.[4] This means that a thesis of inter-world supervenience commits us to the existence of necessary entailments from A- to B-properties.

Kim (1984: 71) shows that if the property sets are assumed to be closed under negation and infinite disjunction, the strict implication between (maximal) supervening and (maximal) subvening properties can be strengthened to a necessary equivalence, taking the infinite disjunction of the different supervenience bases in the different worlds. He argues that necessary equivalence falls short of definability, even in principle:

> "The necessary naturalistic coextension of goodness [...] has no such such epistemological status [i.e. we do not "see" or "infer" that a thing is good by seeing that it has these natural properties (the scare quotes are Kim's)]: we know it must exist, if strong supervenience obtains, but may never know "what it is." Nor can such a coextension be expected to provide a definitional basis for the term "good"; in fact, its existence does not suffice even to show the "in principle" definability of "good" in naturalistic terms. For the notion of definition carries certain semantic and epistemological associations, and even if we could identify the underlying naturalistic coextension of goodness we cannot expect these associations to hold for it." (Kim 1984: 75)

This might be thought too strong: supervenience, after all, was meant to capture determination without definability, even in principle.

So we might opt for the following notion of "global supervenience":

Definition 3 (Global supervenience)**.** *A set of properties A globally supervenes upon a set of properties B iff all possible worlds that are B-indiscernible are A-indiscernible.*[5]

[3] Definitions of strong supervenience equivalent to (2) have been given by Kim (1984: 65), Kim (1987: 81) (who also cites an unpublished paper by Brian McLaughlin), Paull and Sider (1992: 834) and Stalnaker (1996: 89).

[4] Kim (1984: 65) and Kim (1987: 81) show that strong supervenience comes to the following, if the subvening set of properties is closed under infinite conjunctions and disjunctions:

$$\Box\forall x\forall F\in A(Fx\rightarrow\exists G\in B(Gx\wedge\Box\forall y(Gy\rightarrow Fy))) \tag{2}$$

We have an interesting special case if A and B together add up to all properties there are – i.e. when we say that moral or modal properties supervene on the non-moral or non-modal ones. Non-morally or non-modally indiscernible worlds are then said to be indiscernible *tout court*.

I think there are reasons not to identify indiscernibles also in the case of other supervenience theses. Think of Davidson's example of the supervenience of semantics on syntax: there is no semantical difference without a syntactic difference; expressions that are syntactically identical should receive the same semantic analysis. Does this mean that there is only one expression? It does not seem to follow: we could still hold, e.g., that homonyms are best taken to be different words.

Global supervenience logically follows from, but does not logically entail, strong supervenience.[6] They are, however, metaphysically equivalent (hold at the same worlds) if either both the supervening properties A and the subvening properties B are all intrinsic (Paull and Sider 1992: 850) or if both A and B may both contain extrinsic properties and are closed under infinitary Boolean truth-functions, identity and quantification (Stalnaker 1996: 104–105).

Contingency

The more general consideration not to settle against the modal criterion of supervenience is that many supervenience theses are taken to hold contingently:

> "Materialism is meant to be a contingent thesis, a merit of our world that not all other worlds share. Two worlds could indeed differ without differing physically, if at least one of them is a world where Materialism is false." (Lewis 1983: 35)
> "...physicalism is not a claim about every possible world, but only a claim about *our* world to the effect that its physical neature exhausts all its nature." (Jackson 1998: 11)

Global supervenience, as we defined, quantifies over all possible worlds: if metaphysical modality is **S5**, as is standardly assumed in discussions of supervenience, it therefore holds necessarily if at all. The easy way out of the problem is to restrict the range of worlds quantified over: To account for this feature, many analyses of e.g. materialism/physicalism – the thesis that everything supervenes on the physical – have characterised it as modal covariance across a restricted range of possible worlds, i.e. in terms of conditional necessity:

> "Among worlds where no natural properties alien to our world are instantiated, no two differ without differing physically; any two such worlds that are exactly alike physically are duplicates." (Lewis 1983: 37)[7]
> "Any world which is a minimal physical duplicate of our world is a duplicate simpliciter of our world, where a minimal physical duplicate is what you get if you 'stop right there'." (Jackson 1998: 12)

---

[5] Essentially this definition is given by Hellman and Thompson (1975), Haugeland (1982: ??), Horgan (1982: ??), Lewis (1983: 29), Kim (1984: 68), Kim (1987: 82) and (Stalnaker 1996: 91). Horgan and Lewis think it is the appropriate notion.

[6] Kim (1984: 69) asserted their equivalence, but Kim (1987: 82) retracted the right-to-left claim in response to criticism by Hellman (1985), Bacon (1986), Petrie (1987), Teller and Tennant. Paull and Sider (1992) criticise the counterexample of Petrie (1987: 121) ($Fa \wedge Ga$ and $Fa \wedge Gb$ in w_1, $\neg Fa \wedge Ga$ and $\neg Fb \wedge \neg Gb$ in w_2: global, but not strong supervenience), on the grounds that he failed to prove global supervenience and that any such attempt would violate intuitively obvious recombination principles governing the space of possible worlds, i.e. that anything has a lonely intrinsic duplicate, at least if the properties in question are intrinsic. Paull and Sider (1992: 840–841), however, give another counterexample, where the supervening property is extrinsic (in fairness to Petrie it has to be noted that all examples he gives are of extrinsic properties). They argue that strong and global supervenience are equivalent if restricted to intrinsic properties (Paull and Sider 1992: 850). This does not make them equivalent in the case of the physicalists' thesis, though, if, as Sider (2003) has argued, consciousness is a maximal property and maximal properties are extrinsic (cf. Sider 2001).

[7] "Alien" means "neither exemplified by some inhabitant of the actual world nor constructable out of such properties". If "actual" is indexical (as it is on Lewis' theory), then "alien" is so two: a w-alien property then is a property not exemplified by anything in w nor constructable out of properties exemplified in w.

As Hawthorne (2002: 112, n. 8) has remarked, the two accounts are not equivalent, at least assuming an indexical reading of 'actual world' in Lewis' criterion. Lewis' but not Jackson's account rules out a scenario where we have two worlds in which no alien properties are exemplified, which are not minimal duplicates of the actual worlds, but which are physical duplicates, though not duplicates simpliciter, of each other. I think this is a point in favour of Jackson's definition: if materialism is to be a contingent thesis about our world, we should not conclude, from the mere fact that materialism fails in some possible world, that it fails in the actual world.

I think that both accounts are mistaken and neither necessary nor sufficient for physicalism.

If physicalism is contingent, there are some worlds in which it is false. *Prima facie*, there is no reason to assume that no such world can be a minimal physical duplicate or that all such worlds must contain alien properties. Both definitions, however, rule out this scenario; hence, they are not necessary for physicalism. Here is why (the following argument is from Leuenberger (2006)): suppose there is a world w which is a minimal physical duplicate of our world, in which no alien properties are exemplified and in which physicalism fails.

According to Lewis' definition, this means that there are two worlds s and t in which no w-alien properties are exemplified and that are physical duplicates but not duplicates simpliciter. If some properties were exemplified in these worlds that are alien with respect to the actual world, then these properties were also w-alien, for w cannot contain 'more' properties than the actual world. Hence, no alien properties are exemplified in s and t and physicalism fails in the actual world.

According to Jackson's definition, failure of physicalism at w means that w has two minimal physical duplicates, s and t, which are not duplicates simpliciter. Because they are minimal duplicates of w which is a minimal duplicate of our world, they are also minimal duplicates of the actual world and physicalism fails at the actual world.

But is it really possible that physicalism fails in a world which is a minimal physical duplicate of our world and does not contain alien properties?

Because they rule out this intuitively possible scenario, both conditions are not necessary for physicalism. Neither are the two conditions sufficient for physicalism: Hawthorne (2002) has argued that physicalists deny the existence of blockers and that such blockers are not ruled out by the two accounts. Blockers are immaterial entities which prevent the emergence of psychological properties from their alleged physical supervenience base. Because possible, but non-actual, blockers are ruled out by the stop-clause in Jackson's account and because they require the exemplification of alien properties to exist, their possible existence does not falsify physicalism on both accounts.

Even if their definitions are inadequate, Jackson and Lewis might have succeeded in giving us some explication of what it means to be a physicalist. Stalnaker, however, denies even this:

> ...what is interesting, and disquieting, about this way of solving the problem [of making supervenience hold contingently] is that the concept of supervenience is no longer what is doing the work of formulating the reductionist thesis in a way that isolates its metaphysical component. On this account, the materialist's global supervenience thesis is this: relative to all possible worlds that have the same total set of properties and relations as our world, the mental globally supervenes on the physical. But this thesis is a trivial consequence of the materialist thesis that was stated without the notion of supervenience: that the set of all basic properties and relations of our world is the set of physical properties and relations. (Stalnaker 1996: 98)

A way out?

Let us go back to the intuitive idea, aptly characterised as "describ[ing] the same reality in a different way, at a different level of abstraction, by carving the same world at different joints" Stalnaker (1996: 87).

Local, intra-world supervenience is too weak: it does not capture extrinsic dependencies and does therefore not secure that the B-nature of something *alone* secures its A-nature. The property of being among the tallest things

intra-worldy supervenes on the individual height some object has, but something's being among the tallest things does not just depend on its height, but also on the heights of his world-mates. It falls short of the condition that "fixing the base properties of an object fixes its supervenient properties" (Kim 1984: 60). It does not support counterfactuals: we cannot say that if something had the subvening properties, it would also have the supervening ones.

Regional, inter-world is equally too weak, for it does not capture relational dependencies: we would like to say that the property of being smaller than the Eiffel Tower supervenes on height, but two things may differ with respect to it, while being of the same height, if only the Eiffel Tower in their respective worlds is of different height. So it equally falls short of the property-fixing requirement.

Should we therefore settle for global supervenience? I think, with Kim (1987: 86) and Kim (1988: 121), that global supervenience is too strong, too coarse-grained a notion to provide an interesting analysis of dependence and determination. Independent of assumptions about the realm of possibilities, it does not imply weak supervenience and does not rule out worlds containing B-indiscernibles that are A-discernible; neither does not specify property-to-property correlations and does not say of any one individual that *its* A-properties depend on or are exemplified in virtue of *its* B-properties.

What moral can we draw from this discussion?

First, which equivalences and non-equivalences we have between logically different notions of supervenience depends on what kinds of properties we include in the subvening and supervening sets and on which recombination principles we accept for the realm of possibilities. If we restrict both weak and strong supervenience to intrinsic properties, for example, and accept the recombination principle that any two individuals in different worlds have intrinsic duplicates which are world-mates, we can even show that weak entails strong supervenience (cf. Moyer (2000: 4) who also cites Blackburn (1985)). For intrinsic properties, as Paull and Sider have shown, strong and global supervenience stand and fall together.

Second, we do not want strong supervenience with its necessary property-to-property correlations, and therefore we cannot include spatiotemporal relational properties into the subvening base of global supervenience.

Third, global supervenience needs to be supplanted with some coordination of the inhabitants of the worlds in question and it is not clear how this can be achieved: there does not seem a natural and unproblematic way to say that something ties the respective distribution patterns together.

Fourth, it is not clear how we can achieve asymmetry. Fifth, finally, many supervenience theses are meant to be contingent – and it is far from clear how such contingency can be achieved without making them vacuous.

References

Bacon, John, 1986. "Supervenience, necessary coextension, and reducibility". *Philosophical Studies* 49: 163–179 3

Blackburn, Simon, 1985. "Supervenience Revisited". In Hacking, Ian, editor, *Exercises in Analysis: Essays by Students of Casimir Lewy*. Cambridge: Cambridge University Press. Reprinted in Blackburn (1993: 130–148) 5

Blackburn, Simon, 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press 5

Campbell, Keith, 1981. "The Metaphysicd of Abstract Particulars". In French, Peter A., Uehling, Theodore E., Jr., and Wettstein, Howard K., editors, *Midwest Studies in Philosophy 6: Foundations of Analytic Philosophy*, pp. 477–486. Minneapolis, Minnesota: University of Minnesota Press 1

Davidson, Donald, 1970. "Mental Events". In Foster, Lawrence and Swanson, J. W., editors, *Experience and Theory*, pp. 79–101. Amherst, Massachusetts: University of Massachusetts Press. Reprinted in Davidson (1980: 207–225) 1

Davidson, Donald, 1980. *Essays on Actions and Events*. Oxford: Clarendon Press. 2nd, enl. edition: ? 5

Hare, Richard M., 1952. *The Language of Morals*. Oxford: Oxford University Press 1

Haugeland, John, 1982. "Weak Supervenience". *American Philosophical Quarterly* 19: 93–104 2

Hawthorne, John, 2002. "Blocking Definitions of Materialism". *Philosophical Studies* 110: 103–113 3, 4

Hellman, Geoffrey, 1985. "Determination and Logical Truth". *The Journal of Philosophy* 82: 607–616 3

Hellman, Geoffrey and Thompson, F.W., 1975. "Physicalism: ontology, determination, and reduction". *The Journal of Philosophy* 72: 551–564 2

Horgan, Terence E., 1982. "Supervenience and Microphysics". *Pacific Philosophical Quarterly* 63: 29—43 2

Horgan, Terence E., 1993. "From Supervenience to Superdupervenience: Meeting the Demands of a Material World". *Mind* 102: 555–586 1

Jackson, Frank, 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press 1, 3

Kim, Jaegwon, 1984. "Concepts of Supervenience". *Philosophy and Phenomenological Research* 45: 153–176. Reprinted in Kim (1993b: 53–78) 1, 2, 3, 4

Kim, Jaegwon, 1987. " 'Strong' and 'Global' Supervenience Revisited". *Philosophy and Phenomenological Research* 48: 315–326. Reprinted in Kim (1993b: 79–91) 2, 3, 4

Kim, Jaegwon, 1988. "Explanatory Realism, Causal Realism, and Explanatory Exclusion". In French, Peter A., Uehling, Theodore E., Jr., and Wettstein, Howard K., editors, *Midwest Studies in Philosophy 13: Ethical theory – character and virtue*. Notre Dame, Indiana: University of Notre Dame Press 4

Kim, Jaegwon, 1990. "Supervenience as a Philosophical Concept". *Metaphilosophy* 21: 1–27. Reprinted in Kim (1993b: 131–160) 1

Kim, Jaegwon, 1993a. "Postscripts on Supervenience". In Kim (1993b), pp. 161–171 1

Kim, Jaegwon, 1993b. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press 5, 6

Leuenberger, Stephan, 2006. *Ceteris Absentibus Physicalism*. PhD dissertation, Departement of Philosophy, Princeton, New Jersey 4

Lewis, David K., 1983. "New Work for a Theory of Universals". *Australasian Journal of Philosophy* 61: 343–377. Reprinted in Lewis (1999: 8–55) 2, 3

Lewis, David K., 1999. *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press 6

McLaughlin, Brian P., 1992. "The rise and fall of British emergentism". In Beckermann, Ansgar, Flohr, Hans, and Kim, Jaegwon, editors, *Emergence or Reduction?: Prospects for Nonreductive Physicalism*, pp. 49–93. Berlin: Walter de Gruyter 1

Moore, George Edward, 1922. "The Conception of Intrinsic Value". In *Philosophical Studies*, pp. 253–275. London: Routledge and Kegan Paul, Ltd. 1

Moyer, Mark, 2000. "Weak and Global Supervenience: Functional Bark *and* Metaphysical Bite?" Unpublished manuscript, dated 6/2/00 5

Paull, Cranston and Sider, Theodore, 1992. "In defense of global supervenience". *Philosophy and Phenomenological Research* 52: 833–854 2, 3

Petrie, Bradford, 1987. "Global supervenience and reduction". *Philosophy and Phenomenological Research* 48: 119–130 3

Sider, Theodore, 2001. "Maximality and Intrinsic Properties". *Philosophy and Phenomenological Research* 63: 357–362 3

Sider, Theodore, 2003. "Maximality and Microphysical Supervenience". *Philosophy and Phenomenological Research* 66: 139–149 3

Stalnaker, Robert C., 1996. "Varieties of Supervenience". In Tomberlin, James E., editor, *Philosophical Perspectives 10: Metaphysics*, pp. 221–241. Oxford: Basil Blackwell Publishers. Reprinted in Stalnaker (2003: 86–108)] 1, 2, 3, 4

Stalnaker, Robert C., editor, 2003. *Ways A World Might Be: Metaphysical and anti-metaphysical essays*. Oxford: Clarendon Press 6