Australasian Journal of Philosophy
Vol. 72, No. 3; September 1994

PRINCIPLES OF SUPERVENIENCE*

Luc Bovens

The doctrine of supervenience is an attempt to characterize the relationship between families of properties, such as moral and natural properties, mental and physical properties or macro-properties and micro-properties. The literature contains many related, yet by no means equivalent, principles of supervenience. Kim recommends that we 'adopt an attitude of "Let one hundred supervenience concepts bloom!"', since '[each] may have its own sphere of application, serving as a useful tool for formulating and evaluating philosophical doctrines of interest' [3, p.23]. This attitude can be fruitful only if we have a clear understanding of the similarities, differences and the logical connections between alternative principles. Nonetheless, some discussions of supervenience display a curious lack of precision in failing to draw the necessary distinctions. Without claiming to be exhaustive, I will present a categorization of a set of supervenience principles and lay out the logical connections between them.[1] Some isolated pieces of this puzzle have been put together by Kim [2], Klagge [4] and McFetridge [5]. I will sketch the current state of the puzzle and fill in the remaining parts in a systematic fashion. In particular, I will (a) respond to McFetridge's appeal that the logical status of a special set of supervenience principles, viz. of the set of principles of determination[2] be examined [5, pp.257-258], (b) determine the role of complementation and comprehension principles with respect to alternative supervenience principles, (c) provide a solution to a problem first raised by Klagge [4, p.378] involving a curious lack of entailment between supervenience principles.

Let us start with three common slogans that are meant to capture the supervenience relation:

(a) 'Same base-properties, same supervening properties';
(b) 'Different supervening properties, different base-properties';
(c) 'For each supervening property there is some determining base-property'.

Let us try to refine these slogans. First, each slogan is meant to express a necessary truth about the relationship between families of properties and should thus be prefaced by

* This research was partly supported by the Edelstein Center for Philosophy of Science in the Hebrew University of Jerusalem. I am grateful to C.A. Anderson, George Bealer, Carol Cleland, Dalia Drai, Kevin Falvey, James Hawthorne, Paul Moriarty, Chris Shields as well as to anonymous referees of this Journal for their comments on earlier drafts of this paper. I have also benefitted from a presentation of the material at the 1993 ASL Spring Meeting. Proofs of the theorems can be obtained from the author upon request (UC at Boulder, Dept. of Philosophy, CB 232, Boulder CO 80309, USA).

[1] I will not take up the question of diagonal closure [see 1]. This question has some bearing on the present discussion yet can be safely isolated from it.

[2] Or, in McFetridge's terminology, of the set of principles of dependence [5, p.258].

'Necessarily, . . .'. The type of necessity (analytical, metaphysical, . . .) that is proposed will vary relative to one's views about the role of philosophical analysis and the particular sphere of application of the supervenience doctrine. This is of no concern to our present inquiry. Second, each slogan lends itself to three readings. The relationship between supervening properties and base-properties may hold (1) across objects; (2) in a particular subjunctive mode; (3) both across objects *and* in a particular subjunctive mode. For slogan (a) we can thus spell out the following three supervenience principles:

(a_1) 'Necessarily, if an object has the same base-properties as some other object, then those two objects also have the same supervening properties';
(a_2) 'Necessarily, if an object had all of its base-properties held fixed, then it must also retain all of its supervening properties';
(a_3) 'Necessarily, if an object had the same base-properties as some other object, then those two objects must also have the same supervening properties.'

Slogan (b) allows for similar distinctions:

(b_1) 'Necessarily, if an object has different supervening properties than some other object, then those two objects also have different base-properties';
(b_2) 'Necessarily, if an object had different supervening properties, then it must also have different base-properties';
(b_3) 'Necessarily, if an object had different supervening properties than some other object, then those two objects must also have different base-properties.'

For slogan (c) we may introduce similar distinctions in analyzing the expression 'a determining property':

'Necessarily, if an object has a supervening property, then it has a determining base-property',
(c_1) i.e., a base-property which is such that anything that has this base-property also has the supervening property;
(c_2) i.e., a base-property which is such that, if the object had this base-property held fixed, then it must also retain the supervening property;
(c_3) i.e., a base-property which is such that, if anything had this base-property, then it must also have the supervening property.[3]

For a formal presentation of supervenience principles Kim and Klagge restrict themselves to standard modal logic. Standard modal logic only allows for a presentation of principles with subscript '1' and principles matching slogan (c). Let the variables 'x' and 'y' range over the objects that may have supervening properties, the one-place predi-

[3] The objection may be raised that none of the above analyses can be complete since they do not include that some property X determines some property Y only if X is causally efficacious in bringing about Y. But this requirement is plainly false: it is entirely in accordance with ordinarily-language usage to say that the lack of matching blood-samples determines that John is not Karen's biological father.

cate variable '*S*' range over the set of supervening properties and the one-place predicate variable '*B*' range over the set of base-properties. We can now give a formal presentation:

(a_1) $\Box\forall x\forall y[\forall B(Bx \leftrightarrow By) \rightarrow \forall S(Sx \leftrightarrow Sy)]$
(b_1) $\Box\forall x\forall y[\exists S(Sx \,\&\, {\sim}Sy) \rightarrow \exists B(Bx \,\&\, {\sim}By)]$
(c_1) $\Box\forall x\forall S[Sx \rightarrow \exists B(Bx \,\&\, \forall y(By \rightarrow Sy))]$
(c_2) $\Box\forall x\forall S[Sx \rightarrow \exists B(Bx \,\&\, \Box(Bx \rightarrow Sx))]$
(c_3) $\Box\forall x\forall S[Sx \rightarrow \exists B(Bx \,\&\, \Box\forall y(By \rightarrow Sy))]$

By contrast, McFetridge expresses the three supervenience principles matching slogan (a) by quantifying over possible worlds. Let the variables 'v' and 'w' range over possible worlds. Let the two-place predicate variable 'S' range over the set of relations which hold between an object x and a possible world v when x has a supervening property *S* in v. Let the two-place predicate variable 'B' (as well as 'B`', '$B^{\#}$' and 'B^{*}' in subsequent principles) range over the set of relations which hold between an object x and a possible world v when x has a base-property *B* in v. Then (a_1) can be formally expressed without the necessity operator by universally quantifying over a fixed possible-world variable. As to (a_2) and (a_3), to say that the relationship between supervening properties and base-properties holds in this particular subjunctive mode is to say that it holds across possible worlds. Hence, both principles can be formally expressed by quantifying over paired possible-world variables. Our other principles can be formalized in a similar vein. As a mnemonic, let us call all slogan-(a) principles 'COINcidence principles', all slogan-(b) principles 'DIFFerence principles' and all slogan-(c) principles 'DETermination principles'. Within each category I will distinguish between a cross-object principle (. . . ^{co}), a cross-world principle (. . . ^{cw}) and a cross-object/cross-world principle (. . . $^{co/cw}$):

(a_1) or $COIN^{co}$: $\forall v\forall x\forall y[\forall B(Bxv \leftrightarrow Byv) \rightarrow \forall S(Sxv \leftrightarrow Syv)]$
(a2) or COINcw: $\forall v\forall w\forall x[\forall B(Bxv \leftrightarrow Bxw) \rightarrow \forall S(Sxv \leftrightarrow Sxw)]$
(a3) or COINco/cw: $\forall v\forall w\forall x\forall y[\forall B(Bxv \leftrightarrow Byw) \rightarrow \forall S(Sxv \leftrightarrow Syw)]$

(b1) or DIFFco: $\forall v\forall x\forall y[\exists S(Sxv \,\&\, {\sim}Syv) \rightarrow \exists B(Bxv \,\&\, {\sim}Byv)]$
(b2) or DIFFcw: $\forall v\forall w\forall x[\exists S(Sxv \,\&\, {\sim}Sxw) \rightarrow \exists B(Bxv \,\&\, {\sim}Bxw)]$
(b3) or DIFFco/cw: $\forall v\forall w\forall x\forall y[\exists S(Sxv \,\&\, {\sim}Syw) \rightarrow \exists B(Bxv \,\&\, {\sim}Byw)]$

(c1) or DETco: $\forall v\forall x\forall S[Sxv \rightarrow \exists B(Bxv \,\&\, \forall y(Byv \rightarrow Syv))]$
(c2) or DETcw: $\forall v\forall x\forall S[Sxv \rightarrow \exists B(Bxv \,\&\, \forall w(Bxw \rightarrow Sxw))]$
(c3) or DETco/cw: $\forall v\forall x\forall S[Sxv \rightarrow \exists B(Bxv \,\&\, \forall w\forall y(Byw \rightarrow Syw))]$

What are the logical interrelations within this matrix of supervenience principles? Let us first consider COIN/DIFF/DET relations of entailment (or lack of entailment). Kim shows that, on the assumption that supervening and base-properties are 'closed under the usual Boolean property-forming operations, complementation, conjunction, and disjunction' [2, p.158],

$DET^{co} \leftrightarrow COIN^{co}$ [4]

McFetridge establishes the following result. Consider the instantiation of $COIN^{cw}$ for the actual world @:

$$COIN^{cw}{}_{@}: \forall w \forall x[\forall B(Bx@ \leftrightarrow Bxw) \rightarrow \forall S(Sx@ \leftrightarrow Sxw)]$$

The maximal base-property for some object x in @ is the base-property which is coextensive with having all and only x's base-properties in @. Consider the principle which postulates the existence of a maximal base-property for any object x in @:

$$MAXBP_{@}: \forall x \exists B[(Bx@ \ \& \ \forall w(Bxw \rightarrow \forall B`(B`x@ \leftrightarrow B`xw))]$$

Then the conjunction of $COIN^{cw}{}_{@}$ and $MAXBP_{@}$ is *inconsistent* with:

$$*: \exists x \exists S[Sx@ \ \& \ {\sim}\exists B(Bx@ \ \& \ \forall w(Bxw \rightarrow Sxw)]$$

Hence, the conjunction of $COIN^{cw}{}_{@}$ and $MAXBP_{@}$ entails the negation of *, i.e. entails the instantiation of DET^{cw} for @:

$$DET^{cw}{}_{@}: \forall x \forall S[Sx@ \rightarrow \exists B(Bx@ \ \& \ \forall w(Bxw \rightarrow Sxw))]$$

Let us drop Kim's general assumption of closure to determine what entailments are contingent upon particular features of this assumption. First, the following entailments hold, independently of the assumption of closure, for cross-object, cross-world and cross-object/cross-world supervenience:

1. $DET^{..} \rightarrow DIFF^{..}$

2. $DIFF^{..} \rightarrow COIN^{..}$

Generalizing McFetridge's result we learn that the conjunction of $COIN^{cw}$ and the generalization of $MAXBP_{@}$, i.e.,

$$MAXBP: \forall v \forall x \exists B[(Bxv \ \& \ \forall w(Bxw \rightarrow \forall B`(B`xv \leftrightarrow B`xw))]$$

entails DET^{cw}. MAXBP is a comprehension principle in second-order logic which is standardly expressed in the more parsimonious format:

$$COMPR^{cw}: \forall v \forall x \exists B \forall w[Bxw \leftrightarrow \forall B`(B`xv \leftrightarrow B`xw)]$$

We can now state the following theorem:

[4] Or, in Kim's own terminology, definition 1 of weak supervenience $\leftrightarrow$ definition 2 of weak supervenience [2, pp.163-164].

3. $(COIN^{cw} \& COMPR^{cw}) \rightarrow DET^{cw}$

For the cross-object principles, we need to adapt the comprehension principle such that it permits the identification of maximal base-properties across objects rather than across worlds:

$$COMPR^{co}: \forall v \forall x \exists B \forall y[Byv \leftrightarrow \forall B`(B`xv \leftrightarrow B`yv)]$$

The following theorem holds:

4. $(COIN^{co} \& COMPR^{co}) \rightarrow DET^{co}$

Similarly,

5. $(COIN^{co/cw} \& COMPR^{co/cw}) \rightarrow DET^{co/cw}$

with the stronger comprehension principle permitting the identification of maximal base-properties across objects and across worlds:

$$COMPR^{co/cw}: \forall v \forall x \exists B \forall w \forall y[Byw \leftrightarrow \forall B`(B`xv \leftrightarrow B`yw)]$$

Let us now turn to COIN·· and DIFF··. Within each set of principles . . . ^{co}, . . . ^{cw} and . . . $^{co/cw}$, COIN·· entails DIFF·· on the assumption that the base-properties are closed under negation. This assumption can be expressed in complementation principles. The scope of the complementation principles can be more narrow for cross-object and cross-world principles than for cross-object/cross-world principles:

$$COMPL^{co}: \forall v \forall B^{\#} \exists B^{*} \forall x(B^{\#}xv \leftrightarrow \sim B^{*}xv)$$
$$COMPL^{cw}: \forall x \forall B^{\#} \exists B^{*} \forall v(B^{\#}xv \leftrightarrow \sim B^{*}xv)$$
$$COMPL^{co/cw}: \forall B^{\#} \exists B^{*} \forall x \forall v(B^{\#}xv \leftrightarrow \sim B^{*}xv)$$

This brings us to the following theorems:

6. $(COIN^{co} \& COMPL^{co}) \rightarrow DIFF^{co}$

7. $(COIN^{cw} \& COMPL^{cw}) \rightarrow DIFF^{cw}$

8. $(COIN^{co/cw} \& COMPL^{co/cw}) \rightarrow DIFF^{co/cw}$

Figure 1 summarizes the above results:

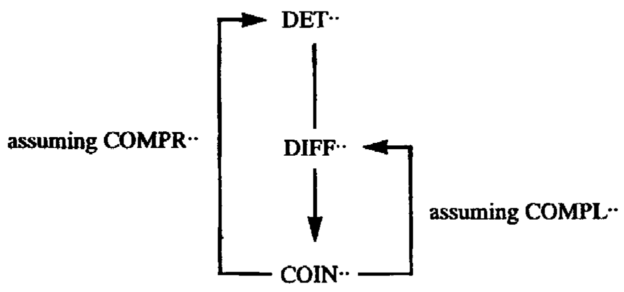

Figure 1

Let us now turn to $(\ldots^{co}/\ldots^{cw}/\ldots^{co/cw})$ relations of entailment (or lack of entailment). Kim notices that $DET^{co/cw}$ entails DET^{co}.[5] Klagge claims that $DET^{co/cw}$ entails DET^{cw}.[6] McFetridge shows that $COIN^{co/cw}$ entails $COIN^{co}$ and $COIN^{cw}$.[7] I may add to these results that $DIFF^{co/cw}$ entails $DIFF^{co}$ and $DIFF^{cw}$. We may list these results in the following theorems:

9. $DET^{co/cw} \rightarrow (DET^{co} \ \& \ DET^{cw})$

10. $DIFF^{co/cw} \rightarrow (DIFF^{co} \ \& \ DIFF^{cw})$

11. $COIN^{co/cw} \rightarrow (COIN^{co} \ \& \ COIN^{cw})$

What about the converse of these entailments? Klagge observes that $(COIN^{co} \ \& \ DET^{cw})$ does not entail $DET^{co/cw}$.[8] The problem is not just that we need to boot up $COIN^{co}$ to DET^{co} by means of $COMPR^{co}$: this would not be sufficient for the entailment to go through. The difficulty lays deeper: suppose that there are two worlds v^{o} and w^{o}. V^{o} contains just object a and w^{o} contains just object b. Let the objects have the same base-properties yet different supervening properties. Then $COIN^{co}$ and DET^{cw} are (trivially) satisfied yet $DET^{co/cw}$ is violated. Klagge writes that he cannot see any condition, short of $DET^{co/cw}$ itself, to conjoin to $(COIN^{co} \ \& \ DET^{cw})$ such that together they entail $DET^{co/cw}$.

Klagge's observation of non-entailment can be generalized. The existence of non-identical objects with cross-world identical base-properties and non-identical supervening properties blocks the entailment relation from the cross-world principle conjoined with the cross-object principle to the cross-object/cross-world principle for COIN, DIFF

[5] Or, in Klagge's own terminology, strong supervenience entails definition 2 of weak supervenience [2, p.166].

[6] Or, in Kim's own terminology, strong synchronic supervenience (SSS) entails the necessary connection condition (NC) [4, p.376].

[7] Or, in McFetridge's own terminology, XYWW' entails XXWW' and XYWW [5, p 257].

[8] Or, in Klagge's own terminology, that weak synchronic supervenience (WSS) conjoined with the necessary connection condition (NC), does not entail strong synchronic supervenience (SSS) [4, p.378 and p.380).

as well as DET. Klagge's observation then follows from the non-entailment relation for DET, conjoined with the claim — which is based on theorems 1,2, and 4 — of COIN^{co} being a weaker principle than DET^{co}.

Does there exist any condition, short of . . . $^{co/cw}$ itself, that can be conjoined to (. . . ^{co} & . . . ^{cw}), such that together they will entail . . . $^{co/cw}$ for either DET, DIFF or COIN? Let us first consider COIN. The entailment can be warranted by stipulating that if there exist objects x in v and y in w with cross-world identical base-properties, then there must be some single world z in which these objects have the same base-properties as in their original worlds. The application of COIN^{cw} to x in v and z, of COIN^{co} to x and y in z and of COIN^{cw} to y in w and z warrants that $\mathrm{COIN}^{co/cw}$ will be respected for x in v and y in w. This stipulation can be expressed in the following mixed-world condition:

MW: $\forall v \forall w \forall x \forall y [\forall B(Bxv \leftrightarrow Byw) \rightarrow \exists z \forall B((Bxv \leftrightarrow Bxz) \ \& \ (Byw \leftrightarrow Byz))]$

This brings us to the following theorem:

12. $(\mathrm{COIN}^{co} \ \& \ \mathrm{COIN}^{cw} \ \& \ \mathrm{MW}) \rightarrow \mathrm{COIN}^{co/cw}$

and hence, by theorems 2, 8 and 12:

13. $(\mathrm{DIFF}^{co} \ \& \ \mathrm{DIFF}^{cw} \ \& \ \mathrm{MW} \ \& \ \mathrm{COMPL}^{co/cw}) \rightarrow \mathrm{DIFF}^{co/cw}$

and by theorems 1, 2, 5 and 12:

14. $(\mathrm{DET}^{co} \ \& \ \mathrm{DET}^{cw} \ \& \ \mathrm{MW} \ \& \ \mathrm{COMPR}^{co/cw}) \rightarrow \mathrm{DET}^{co/cw}$

Figure 2 summarizes the above results:

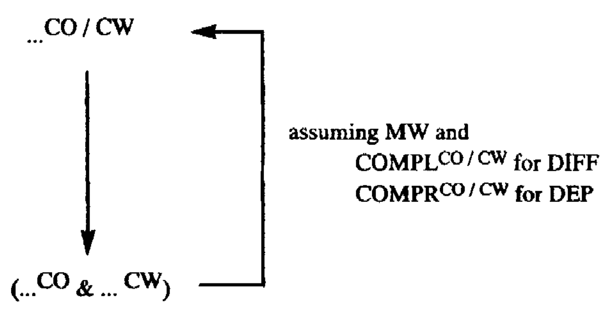

Figure 2

We can now solve Klagge's problem, i.e. what conditions conjoined with (COIN^{co} & DET^{cw}) entail $\mathrm{DET}^{co/cw}$? Assuming COMPR^{co}, COIN^{co} entails DET^{co} (theorem 4) and assuming (MW & $\mathrm{COMPR}^{co/cw}$), (DET^{co} & DET^{cw}) entails $\mathrm{DET}^{co/cw}$ (theorem 14). These conditions can be simplified in the light of the following theorem:

15. $COMPR^{co/cw} \rightarrow (COMPR^{co} \& COMPR^{cw})$

Hence, Klagge's problem can be solved by conjoining ($COIN^{co}$ & DET^{cw}) with the condition (MW & $COMPR^{co/cw}$) to entail $DET^{co/cw}$.

The above results provide the groundwork for raising further philosophical questions about supervenience. To determine what supervenience relation is fitting within a particular sphere of application the plausibility of the intermediary principles needs to be addressed. For instance, Van Cleve offers several arguments to the effect that an indiscriminate acceptance of the complementation principle makes supervenience into a trivial relationship: what is needed are independent grounds to the effect that the complement of some base-property — say, some physical property — is itself a physical property [6, pp.229-231]. One might also question a blanket endorsement of comprehension principles considering that they may be trivially satisfied by including some base-property which cannot be adopted across objects within a single possible world — say, haecceities or spatio-temporal location — or which cannot be adopted across possible worlds — say, actuality. MW is controversial in that it excludes base-properties which can be instantiated by at most one object within a possible world. Yet these questions are beyond the scope of this paper.

University of Colorado at Boulder

Received November 1992
Revised August 1993

REFERENCES

1. J. Bacon, 'Supervenience, Necessary Coextension, and Reducibility', *Philosophical Studies* 49 (1986) pp.163-176.
2. J. Kim, 'Concepts of Supervenience', *Philosophy and Phenomenological Research* 45 (1984) pp.153-176.
3. J. Kim, 'Supervenience as a Philosophical Concept', *Metaphilosophy* 21 (1990) pp.1-27.
4. J. Klagge, 'An Alleged Difficulty Concerning Moral Properties', *Mind* 93 (1984) pp.370-380.
5. J.G. McFetridge, 'Supervenience, Realism, Necessity', *The Philosophical Quarterly* 35 (1985) pp.245-258.
6. J. Van Cleve, 'Supervenience and Closure', *Philosophical Studies* 58 (1990) pp.225-238.