Supervenience, reduction, and translation

February 23, 2018

This paper considers the following question: what is the relationship between supervenience and reduction? I investigate this formally, first by introducing a recent argument by Christian List to the effect that one can have supervenience without reduction; then by considering how the notion of Nagelian reduction can be related to the formal apparatus of definability and translation theory; then by showing how, in the context of propositional theories, topological constraints on supervenience serve to enforce reducibility; and finally, how constraints derived from the theory of ultraproducts can enforce reducibility in the context of first-order theories.

1 List on supervenience and reduction

I'll start by giving a brief recapitulation of the apparatus used by List (2018) to analyse supervenience. For List, a *system of ontological levels* is a concrete posetal category wherein each function is surjective. That is, it consists of a class of sets $\mathcal{L}$, equipped with surjective functions between them, such that (i) the class of all such functions is closed under composition, and for each set we include the identity function on that set (hence making this a concrete category); and (ii) for any sets A and B, there is a *unique* function σ mapping A to B (hence making the category posetal).

In order to relate this to issues of reduction, List introduces certain linguistic notions. Formally, he defines a *language L* to be a set (of "sentences"), equipped with a unary "negation" operator and a "consistency" property. Given a language L, List defines the *ontology Ω_L* of L to be the set of all maximal consistent subsets of L (where a consistent set A is *maximal* iff there is no consistent B such that $A \subset B$). Each element of this ontology he calls a *world* for L. For any $\phi \in L$, ϕ is *true at* the world $\omega \in \Omega_L$ iff $\phi \in \omega$; otherwise, it is false at ω. The *propositional content* $[[\phi]]$ of $\phi \in L$ is the set of worlds at

which ϕ is true: that is, $[[\phi]] := \{\omega \in \Omega_L : \phi \in \omega\}$. Then, a *system of descriptive levels* is a system of ontological levels where each level is the ontology Ω_L of some language L.

The notion of reduction is defined as follows. Given some system of descriptive levels, suppose that $\sigma : \Omega_L \to \Omega_{L'}$ is a supervenience mapping. List says that a sentence $\phi' \in L'$ *reduces* to $\phi \in L$ if $[[\phi]] = \sigma^{-1}([[\phi']])$: that is, if ϕ is true at a world ω iff ϕ' is true at $\sigma(\omega)$. And we say that the language L' reduces to L if every $\phi' \in L$ reduces to some $\phi \in L$: that is, if every higher-level sentences is reducible to some lower-level sentence.

Certainly, setting up this kind of association between the apparatus of possible worlds and languages is going to be necessary if we are going to talk about the relationship between supervenience (a map, on this treatment, between ontologies) and reduction (a relationship between languages). But it is not clear to me that this is quite the right way to set things up.

First, it seems a mistake to use maximal consistent sets of sentences to represent the ontology associated to a language. This is not because I have a problem with identifying worlds with sets of sentences—we're just doing mathematical modelling, after all, so "identification" isn't really doing anything more than asserting a one-to-one correspondence. Rather, the problem is that there are intuitive reasons to think that, in general, there are more worlds (of a given language's ontology) than there are maximal consistent sets of sentences of that language. Specifically, consider the class of models of a first-order language L. It seems plausible to suppose that there could be a world just like each model—one with the same number of individuals as there are elements in the model's domain, and with properties and relations distributed over those individuals just as the extensions of the predicates are distributed over the model. This suggests that we can take such models to represent worlds, with non-isomorphic models representing distinct worlds.[1] But in first-order logic, there are non-isomorphic models which are elementarily equivalent, i.e., which satisfy all the same sentences. So a maximal consistent set of sentences will, in general, be satisfied by several non-isomorphic models, and so by several distinct worlds.

For this reason, I suggest that a friendly amendment to List's proposal is that we take the ontology associated with a language to be the class of models for that language, rather than the set of maximal consistent sets of sentences of that language. This will require us to say a little more about the nature of the language: for the purposes of this essay, I will assume that we are working with finitary first-order languages, each

---

[1]Whether isomorphic models represent distinct worlds or not is more controversial (it is closely related to the question of whether or not to believe in haecceities, i.e., primitive trans-world identities for individuals).

with a particular nonlogical vocabulary. Note that we are including the possibility that the language is a propositional language (regarding propositional constants as nullary predicates).

That said, it does seem that we don't want to represent the ontology associated to a given level by *all* models of the associated language. After all, if the language contains predicates R for "red all over" and G for "green all over", then there will be a model containing objects satisfying both R and G. That is, the worlds represented by the class of all models is intuitively the class of logically possible worlds, but that is typically not the species of possibility that we are interested in for the purposes of supervenience. Standardly, we want to consider supervenience relative to metaphysical possibility, or relative to something more restrictive still (e.g. physical possibility). In List's approach, one could code this up into the consistency property, so that only worlds corresponding to metaphysically or physically consistent sets of sentences are permitted (where "physical inconsistency" would mean, roughly, consistency with the physical laws).

In the friendly amendment I'm suggesting here, a more natural way to do things would be to confine attention to the models *of some theory*, where that theory consists of all and only the sentences expressing things which are necessary—necessary, that is, relative to the standard of possibility we seek to capture. Thus, if we are thinking about metaphysical possibility, then we take the theory to consist of all the metaphysically necessary L-sentences (e.g. $\forall x(\neg Rx \wedge Gx)$); if we are thinking about physical possibility, then we take the theory to consist of all the L-sentences expressing physical laws; and so on and so forth. So every descriptive level is associated with a given theory, and the supervenience maps between levels are maps between the sets of models of the associated theories. We follow List in imposing the requirements that the sets of models, equipped with the supervenience maps, constitute a posetal category. Rather than requiring that these maps be surjective, however, it will be more helpful to reflect further on the relationship between models and the possible worlds they represent.

The relevant observation is that consistency with the laws at its own level is only a necessary condition for a model to represent a possible world, not a sufficient one: it is natural to think that the sense of possibility relevant for supervenience is something stronger. Just because a certain distribution of thermodynamical properties is possible according to the laws of thermodynamics, for instance, does not mean that that distribution is possible *tout court*; in particular, such a distribution may not be one that could actually be generated by any statistical-mechanically possible configuration of the microstructure. Thus, in the event that two descriptive levels, associated with theories

T_1 and T_2, have a supervenience map $\sigma : \text{Mod}(T_2) \to \text{Mod}(T_1)$, then we should take the ontology of T_1 (i.e., the possible worlds at the T_1-level) to include only those models of T_1 that lie in the image of the supervenience mapping: i.e., that $\Omega_1 \subseteq \sigma[\text{Mod}(T_2)]$. And of course, for any other descriptive level upon which the T_1-level supervenes, the same requirement will hold.

By taking these to be the only such requirements, we may define the ontology associated with a given level as follows: suppose that the T-level supervenes upon all and only the T_i-levels, for i in some index set I, with the supervenience mappings $\sigma_i : \text{Mod}(T_i) \to \text{Mod}(T)$. Then we let the ontology for the T-level be $\Omega := \bigcap_{i \in I} \sigma_i[\text{Mod}(T_i)]$. It follows that if the T_1-level supervenes upon the T_2-level, then the supervenience map is surjective on ontologies: that is, for every $\mathcal{M}_1 \in \Omega_1$, there is some $\mathcal{M}_2 \in \Omega_2$ such that $\sigma(\mathcal{M}_2) = \mathcal{M}_1$.

Finally, if the T_1-level supervenes upon the T_2-level, with supervenience map σ, I'll say that an L_1-sentence ϕ_1 *List-reduces* to the L_2-sentence ϕ_2 if $[[\phi_2]] = \sigma^{-1}([[\phi_1]])$: that is, if for every $\mathcal{M} \in \Omega_2$, $\mathcal{M} \models \phi_2$ iff $\sigma(\mathcal{M}) \models \phi_1$. And T_1 *List-reduces* to T_2 if every L_1-sentence ϕ_1 is reducible to some L_2-sentence ϕ_2.

List observes that, within his framework, supervenience does not entail reducibility. In the version of that framework developed here, this is encoded by the following result.

Theorem 1. *Not every system of descriptive levels is such that if the T_1-level supervenes on the T_2-level, then T_1 is List-reducible to T_2.*

List appeals to general combinatorial considerations for a proof. However, it will be illuminating (and helpful for the sequel) to construct an explicit counterexample to the hypothesis that supervenience entails reduction.[2]

Proof. Let L_1 be the propositional language whose only sentence-letter is F, and let L_2 be the propositional language with sentence-letters $\{P_0, P_1, \dots\}$. Let $T_1 = T_2 = \varnothing$. $\text{Mod}(T_1)$ only contains two worlds: set $\mathcal{M}_F(F) = \top$, and $\mathcal{M}_{\neg F}(F) = \bot$. Let $\mathcal{M}$ be the T_2-model such that $\mathcal{M}(P_i) = \top$ for every P_i. Define $\sigma : \text{Mod}(T_2) \to \text{Mod}(T_1)$ as follows: for any $\mathcal{A} \in \text{Mod}(T_2)$,

$$\sigma(\mathcal{A}) := \begin{cases} \mathcal{M}_F & \text{if } \mathcal{A} = \mathcal{M} \\ \mathcal{M}_{\neg F} & \text{otherwise} \end{cases} \tag{1}$$

Since the T_2-level is the lowest, $\Omega_2 = \text{Mod}(T_2)$; hence $\Omega_1 = \sigma[\Omega_2] = \text{Mod}(T_1)$.

---

[2]This example is inspired by a construction in Halvorson (2012).

Now observe that $\sigma^{-1}([[F]]) = \{\mathcal{M}\}$. But $\{\mathcal{M}\}$ is not a definable subset of Ω_2: that is, there is no sentence $\phi \in L_2$ such that $[[\phi]] = \{\mathcal{M}\}$.[3] So the sentence $F \in L_1$ is not reducible to any sentence in L_2; thus, T_1 is not reducible to T_2. $\qquad\square$

Now, one natural (and I think appropriate) response to this argument is that it is unsurprising. After all, we have imposed almost no constraints on the supervenience map. A supervenience map between two ontological levels is just a map from the worlds of one level to the worlds of the other, with no constraints on how that map must relate to the structure present in those worlds. So for the rest of this essay, I would like to consider some ways in which we could impose such constraints on the maps, and what consequences those constraints have for reducibility.

2 Definition and reduction

Before going further, I want to pause to make contact with some more fully-fledged ways of thinking about reduction. This will involve some of the apparatus of definability theory from logic.[4] An *explicit definition* of R in terms of Σ is a formula of the form

$$\delta_R = \forall x_1 \ldots \forall x_n (Rx_1 \ldots x_n \leftrightarrow \tau_R(x_1, \ldots, x_n)) \tag{2}$$

where τ_R is a Σ-formula. Given a Σ-theory T and Σ^+-theory T^+, where $\Sigma \subset \Sigma^+$, T^+ is a *definitional extension* of T if T^+ is logically equivalent to the theory

$$T \cup \{\delta_R : R \in \Sigma^+ \setminus \Sigma\} \tag{3}$$

where for each R, δ_R is an explicit definition of R in terms of Σ. More generally, any Σ^+-theory T^+ explicitly defines a symbol $R \in \Sigma^+ \setminus \Sigma$ in terms of Σ if it entails an explicit definition δ_R of R in terms of Σ: that is, if $T^+ \vDash \delta_R$ for some δ_R of the form (2).

Now, this may be compared to Nagel (1979)'s definition of reduction:

> [. . .] a reduction is effected when the experimental laws of the secondary science (and if it has an adequate theory, its theory as well) are shown to be the logical consequences of the theoretical assumptions (inclusive of the

---

[3]Proof: suppose for reductio that $[[\phi]] = \{\mathcal{M}\}$. Since ϕ is a finite sentence, not every sentence-letter can occur in it. So suppose P_i does not occur in ϕ. Then since $\mathcal{M} \vDash \phi$, it must be the case that $\mathcal{M}' \vDash \phi$, where $\mathcal{M}'$ is just like $\mathcal{M}$ save that $\mathcal{M}'(P_i) = \bot$ (as the truth-value of a sentence in propositional logic is dependent only on the truth-values of the sentence-letters occurring in it). But then $\mathcal{M}' \in [[\phi]]$, although $\mathcal{M}' \neq \mathcal{M}$, so we have a contradiction. QED.

[4]See (Hodges, 1997, §2.6) for further details.

coordinating definitions) of the primary science. [. . .] when the laws of the secondary science do contain some term 'A' that is absent from the theoretical assumptions of the primary science, there are two necessary formal conditions for the reduction of the former to the latter: (1) Assumptions of some kind must be introduced which postulate suitable relations between whatever is signified by 'A' and traits represented by terms already present in the primary science. [. . .] (2) With the help of these additional assumptions, all the laws of the secondary science, including those containing the term 'A', must be logically derivable from the theoretical premises and their associated coordinating definitions in the primary discipline.[5]

As it is often put: a theory T_1 is reducible to a theory T_2 just in case T_1 can be derived from T_2 together with so-called *bridge laws*. Now, as has been widely discussed,[6] if we place no restrictions on the content of the bridge laws then reduction becomes trivialised: any theory T_1 may be reduced to any other theory T_2, by taking as the bridge laws either T_1 itself, or any inconsistent set of statements. To avoid this problem, we put two restrictions on the bridge laws. The first is that the bridge laws do indeed serve to bridge the gap between T_1 and T_2: formally, we require that T_2 plus the bridge laws explicitly defines every symbol $R \in \Sigma_1 \setminus \Sigma_2$ (where Σ_1 and Σ_2 are, respectively, the vocabularies of T_1 and T_2). The second is that the bridge laws not sneak in "extra content" beyond this bridging function: formally, we require that T_2 plus the bridge laws be a *conservative extension* of T_2. (Recall that a Σ^+-theory T^+ is a conservative extension of a given Σ-theory T, where $\Sigma \subseteq \Sigma^+$, if T^+ has no new Σ-consequences relative to T: that is, if for every Σ-sentence ϕ, $T^+ \vDash \phi$ entails that $T \vDash \phi$.)

But if these two conditions are satisfied, then T_2 together with the bridge laws is a definitional extension of T_2.[7] Thus, to within logical equivalence, we may suppose that the bridge laws *are* explicit definitions of the novel terms (i.e., of the symbols in $\Sigma_1 \setminus \Sigma_2$).[8] Note that this agrees with Schaffner's requirement that bridge laws be "reduction functions" (statements that a certain term of T_1 and a certain term of T_2 are coextensional)—provided, that is, that we take the reduction functions to relate a simple term of T_1 to a (possibly) complex term of T_2.[9] Thus, we are led to posit the following

---

[5](Nagel, 1979, pp. 352–354)

[6]See, for instance, Dizadji-Bahmani et al. (2010) and references therein.

[7](Hodges, 1997, pp. 53–54).

[8]Note that all we are supposing here is that every bridge law has the syntactic form of a definition (i.e., of a biconditional of the form (2)). This is neutral on the question of what the *content* of the bridge laws is: for instance, whether that of a set of definitions, a set of conventional stipulations, or a set of factual assertions (Nagel, 1979, pp. 354–355).

[9]This identification of Nagelian reduction with definitional extension is also made by Butterfield (2011a).

relation: a theory T_1 *Nagel-reduces* to T_2 if there is some definitional extension T_2^+ of T_2 such that $T_2^+ \vDash T_1$.[10]

We now want to think about how all this relates to the apparatus of supervenience mappings developed in the previous section. For these purposes, it is helpful to work with the notions of interpretation and translation rather than definition—though as we shall see, these are closely intertwined.[11] So, define an *interpretation* from one language L_1 to another language L_2 to be a map $\tau : \mathrm{Form}(L_1) \to \mathrm{Form}(L_2)$ which commutes with the logical constants (where $\mathrm{Form}(L)$ is the set of formulae of L). And given theories T_1 and T_2, in languages L_1 and L_2 respectively, define a *translation* from T_1 to T_2 to be an interpretation $\tau : L_1 \to L_2$ such that for any $\phi \in \mathrm{Form}(L_1)$, if $T_1 \vDash \phi$, then $T_2 \vDash \tau(\phi)$. This lets us give a more compact characterisation of Nagel-reduction, as follows.

Theorem 2. T_1 *Nagel-reduces to* T_2 *iff there is a translation* $\tau : T_1 \to T_2$ *which restricts to the identity on* $\Sigma_1 \cap \Sigma_2$.

Proof. Before starting the proof proper, we state a quick lemma. Suppose that $\tau : \Sigma_1 \to \Sigma_2$ is an interpretation which restricts to the identity on $\Sigma_1 \cap \Sigma_2$. For any n-ary $R \in \Sigma_1 \setminus \Sigma_2$, let $\delta_R := \forall x_1 \ldots \forall x_n (R x_1 \ldots x_n \leftrightarrow \tau(R)(x_1, \ldots, x_n))$. Then for any $\phi \in \mathrm{Form}(\Sigma_1 \cup \Sigma_2)$, $\{\delta_R : R \in \Sigma_1 \setminus \Sigma_2\} \vDash \phi \leftrightarrow \tau(\phi)$. The lemma can be demonstrated by a straightforward proof by induction on the complexity of formulae.

Now, suppose that T_1 Nagel-reduces to T_2, i.e., that $T_2^+ \vDash T_1$ for some definitional extension T_2^+ of T_2. Then for any $R \in \Sigma_1 \setminus \Sigma_2$, $T_2^+ \vDash \delta_R$ where δ_R is some explicit definition of R, i.e., some formula of the form (2). So let $\tau(R) := \tau_R$ for any such R (where τ_R is the Σ_2-formula occurring on the right-hand-side of the biconditional in δ_R). For any $S \in \Sigma_1 \cap \Sigma_2$, let $\tau(S) = S$. Now extend τ to a map from $\mathrm{Form}(\Sigma_1)$ to $\mathrm{Form}(\Sigma_2)$ by demanding that it commute with the logical constants. So τ is an interpretation from Σ_1 to Σ_2 (which restricts to the identity on $\Sigma_1 \cap \Sigma_2$), and it only remains to show that it is a translation from T_1 to T_2.

So for an arbitrary Σ_1-formula ϕ, suppose $T_1 \vDash \phi$. Then $T_2^+ \vDash \phi$, since $T_2^+ \vDash T_1$. By the lemma, $T_2^+ \vDash \phi \leftrightarrow \tau(\phi)$, and hence $T_2^+ \vDash \tau(\phi)$. But $\tau(\phi)$ is a Σ_2-formula, and T_2^+ is a conservative extension of T_2; so $T_2 \vDash \tau(\phi)$. This suffices to show that τ is a translation, as requested.

Next, suppose that τ is a translation from T_1 to T_2 which restricts to the identity on $\Sigma_1 \cap \Sigma_2$. For every $R \in \Sigma_1 \setminus \Sigma_2$, let $\delta_R := \forall x_1 \ldots \forall x_n (R x_1 \ldots x_n \leftrightarrow \tau(R)(x_1, \ldots, x_n))$.

---

[10]Note that we are only permitting the definition of new relations, and not of any new objects in the domain: in more metaphysical terms, we are considering only reductions of properties and not of individuals. Obviously, this is rather limiting (the individuals of physics are clearly distinct from those of biology!)—but we will keep this restriction, in the interests of simplicity.

[11]cf. Barrett and Halvorson (2016)

Now let $T_2^+ := T_2 \cup \{\delta_R : R \in \Sigma_1 \setminus \Sigma_2\}$. So clearly, T_2^+ is a definitional extension of T_2, and it remains only to show that $T_2^+ \vDash T_1$.

So consider any $\phi \in T_1$. Then $T_2 \vDash \tau(\phi)$, since τ is a translation. So $T_2^+ \vDash \tau(\phi)$. By the lemma, $T_2^+ \vDash \phi \leftrightarrow \tau(\phi)$; so, $T_2^+ \vDash \phi$. This suffices to prove that $T_2^+ \vDash T_1$, and hence that T_1 Nagel-reduces to T_2. □

Let the class of Σ-structures be denoted $\mathrm{Str}(\Sigma)$. Now, given an interpretation $\tau : \Sigma_1 \to \Sigma_2$, the *dual map* of τ is a map $\tau^* : \mathrm{Str}(\Sigma_2) \to \mathrm{Str}(\Sigma_1)$, defined as follows: for any Σ_2-structure $\mathcal{A}$, $|\tau^*(\mathcal{A})| = |\mathcal{A}|$ (where $|\mathcal{M}|$ denotes the domain of the model $\mathcal{M}$); and for any $R \in \Sigma_1$ and $a_1, \dots, a_n \in |\mathcal{A}|$, its extension $R^{\tau^*(\mathcal{A})}$ is the set of all and only those n-tuples which satisfy $\tau(R)$. It is straightforward to show that for any Σ_2-structure $\mathcal{A}$ and Σ_1-sentence ϕ, $\tau^*(\mathcal{A}) \models \phi$ iff $\mathcal{A} \models \tau(\phi)$. As a corollary, if τ is a translation from T_1 to T_2, then $\tau^* : \mathrm{Mod}(T_2) \to \mathrm{Mod}(T_1)$ (i.e., if $\mathcal{A}$ is a model of T_2, then $\mathcal{A}$ is a model of T_1).

Putting this all together, we get the following: if T_1 Nagel-reduces to T_2, then the reduction naturally induces a map from $\mathrm{Mod}(T_2)$ to $\mathrm{Mod}(T_1)$. Moreover, if the translations between a collection of theories constitute a posetal category, then their dual maps will as well: this follows from the fact that given translations $\tau_1 : T_1 \to T_2$ and $\tau_2 : T_2 \to T_3$, $(\tau_2 \circ \tau_1)^* = \tau_1^* \circ \tau_2^*$. So such a collection of translations gives rise to a system of descriptive levels of the kind motivated in the previous section. Thus, Nagel-reduction entails supervenience. Moreover, it entails List-reduction: given any L_1-sentence ϕ_1 and a translation $\tau : T_1 \to T_2$, ϕ_1 List-reduces to $\tau(\phi_2)$. It follows that the example in Theorem 1 also shows that supervenience does not entail Nagel-reducibility.

Incidentally, the dual maps will not typically be surjective (as maps between classes of models)—nor would it be appropriate to require that they were. To see this, observe that the following principle seems plausible: if T_1 reduces to T_2, and if T_0 is a subtheory of T_1 (i.e. if every sentence of T_0 is a sentence of T_1) then T_0 reduces to T_2. But it is an immediate consequence of this that not all dual maps can be surjective. Let T_0 be a proper subtheory of T_1, let τ be the translation from T_1 to T_2, and observe that it is thereby a translation from T_0 to T_2. If τ^* is surjective as a map from $\mathrm{Mod}(T_2)$ to $\mathrm{Mod}(T_1)$, then we are done. If not, then since T_0 is a proper subtheory of T_1, there is at least one model $\mathcal{M}$ of T_0 which is not a model of T_1. Since the image of $\mathrm{Mod}(T_2)$ under τ^* is $\mathrm{Mod}(T_1)$, then $\mathcal{M}$ is *not* in said image of, and so τ^* is not surjective as a map from $\mathrm{Mod}(T_2)$ to $\mathrm{Mod}(T_0)$.[12] As discussed previously, however, we can impose surjectivity of supervenience maps *as maps on possible worlds* by fiat.

---

[12]This poses a problem for Suppes' account of reduction (Suppes, 1957, 1967), given that he effectively presupposes surjectivity of the map from the lower-level theory to the higher-level theory associated with reduction.

With all this in hand, we now return to our main task: what kinds of constraints could be imposed on supervenience maps, in order to guarantee that they are associated with some form of reducibility? First, I will consider the idea that the supervenience maps must preserve *similarity*, as encoded in topological data about the models.

3 Supervenience and continuity

In his (2011b), Butterfield asks us to consider the sequence of functions $g_n : \mathbb{R} \to \mathbb{R}$, for $n \in \mathbb{N}^+ := \mathbb{N} \setminus \{0\}$, where

$$g_n(x) := \begin{cases} -1 & \text{if } x < -\frac{1}{n} \\ nx & \text{if } -\frac{1}{n} \le x \le \frac{1}{n} \\ 1 & \text{if } \frac{1}{n} < x \end{cases} \tag{4}$$

The limit of this sequence (as determined by pointwise convergence) is the function g_∞, where

$$g_\infty(x) := \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } 0 < x \end{cases} \tag{5}$$

As Butterfield notes, for each finite n the function g_n is continuous; yet the limit of the sequence of these individually continuous functions, g_∞, is discontinuous (at 0). Now consider the function $f_\bullet : \mathbb{N}^+ \cup \{\infty\} \to \{0, 1\}$, defined by

$$f_n := \begin{cases} 0 & \text{if } n \in \mathbb{N}^+ \\ 1 & \text{if } n = \infty \end{cases} \tag{6}$$

We can now define a two-level ontological system, in the sense of List, by taking one ontology to be the set $\{g_n\}_{n\in\mathbb{N}^+} \cup \{g_\infty\}$, the other to be the set $\{f_n\}_{n\in\mathbb{N}^+\cup\{\infty\}} = \{0, 1\}$, and the supervenience map to be $\sigma : g_n \mapsto f_n$. Alternatively, we could take the lower-level ontology to consist of all functions $\mathbb{R} \to \mathbb{R}$, the higher-level ontology to consist (again) of the set $\{0, 1\}$, and the supervenience function to be a function mapping continuous functions to 0 and discontinuous functions to 1.

This supervenience map fails to preserve limits. The limit of the sequence of lower-level worlds $g_1, g_2, \ldots$ is g_∞, and $\sigma(g_\infty) = 1$. By contrast, the limit of the sequence $\sigma(g_1), \sigma(g_2), \ldots$ is the limit of the sequence $0, 0, \ldots$, which is of course 0. So it makes a

difference whether we take the limit before or after applying the supervenience map; taking the limit does not commute with supervenience. Thinking of topological structure as giving a standard of similarity among the models,[13] let us say that a supervenience map which fails to commute with limits in this fashion is not *similarity-preserving*.

But now consider again the counterexample constructed in the proof of Theorem 1. We can make this example resembles Butterfield's example of the functions, by defining the models $\mathcal{M}_n$, for $n \in \mathbb{N}$, as follows: $\mathcal{M}_n$ assigns $\top$ to P_0 through P_n, and $\bot$ to all remaining sentence-letters. Then there is an intuitive sense in which the models in the sequence $\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \ldots$ get "more and more similar" to the model $\mathcal{M}$—that is, a sense in which we can regard $\mathcal{M}$ as the limit of the sequence $\mathcal{M}_n$ as n goes to infinity. This intuition receives a precise expression in the concept of the *Stone topology* of a class of propositional models: in that topology, a sequence of models $\mathcal{M}_n$ converges to a model $\mathcal{M}$ if and only if the sequence of truth-values $P^{\mathcal{M}_n}$ converges to the truth-value $P^{\mathcal{M}}$, for every sentence-letter P.[14] However, the supervenience map does not play nice with this convergence structure: just as with Butterfield's example, the supervenient image of the limit of the sequence (i.e., $\mathcal{M}_F$) is not the limit of the supervenient images of the sequence (i.e., $\mathcal{M}_{\neg F}$). In other words, the supervenience map is not similarity-preserving.

This suggests a conjecture: perhaps all such failures of reduction are associated with the failure of the supervenience map to be similarity-preserving? In the propositional case, this turns out to be correct. More specifically, we have the following result.

Theorem 3. *Suppose that T_1 and T_2 are propositional theories. Then there is a translation $\tau : T_1 \to T_2$ iff there is a map $\sigma : \mathrm{Mod}(T_2) \to \mathrm{Mod}(T_1)$ which is continuous in the Stone topology. In the event that this holds, $\sigma = \tau^*$.*

Proof. First, suppose that τ is a translation from T_1 to T_2. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be the Lindenbaum-Tarski algebras of T_1 and T_2.[15] By Stone's representation theorem, any homomorphism $\rho : \mathcal{T}_1 \to \mathcal{T}_2$ is associated with a (Stone-)continuous function $\sigma : \mathrm{Mod}(T_2) \to \mathrm{Mod}(T_1)$ and vice versa. So in one direction, the translation $\tau : T_1 \to T_2$ gives rise to a homomorphism $\rho : \mathcal{T}_1 \to \mathcal{T}_2$, and thence to a continuous function σ, which coincides with the dual map τ^*. And in the other, a continuous function σ gives rise to a homomorphism ρ; but any homomorphism between atomic Boolean algebras

---

[13]cf. Fletcher (2016)

[14]Halvorson (2012)

[15]The Lindenbaum-Tarski algebra of a propositional theory T is a Boolean algebra whose elements are T-equivalent sets of sentences, and whose meet, join and complement are given by (the abstractions of) the conjunction, disjunction and negation operators.

is generated by a map from the atoms of one to those of the other, and so ρ is generated by a translation $\tau : T_1 \to T_2$, which is such that $\tau^* = \sigma$. $\qquad\qquad\square$

4 The first-order case

Let us now turn our attention to the case of first-order theories. It will be helpful to split our enquiry into two parts. First, can we find conditions on a supervenience mapping sufficient for that mapping to be associated with a reduction? Second, can we think of those conditions as imposing a requirement of similarity-preservation?

With regard to the first question, the following result—based on a result by van Benthem and Pearce (1984)—provides an affirmative answer.

Theorem 4. *Suppose that T_1 and T_2 are first-order theories. Then there is a translation $\tau : T_1 \to T_2$ iff there is a map $\sigma : \mathrm{Mod}(T_2) \to \mathrm{Mod}(T_1)$ which preserves isomorphisms, ultraproducts and domains.[16] In the event that this holds, $\sigma = \tau^*$.*

Before turning to the proof, I review the relevant notion of an ultraproduct.[17] First, an *ultrafilter* over a set X is a non-empty set U of subsets of X, such that

- $\varnothing \notin U$

- U is closed under intersection: if $A \in U$ and $B \in U$, then $A \cap B \in U$

- U is closed under supersets: if $A \in U$ and $A \subseteq B$ (for $B \subseteq X$), then $B \in U$

- For every subset A of X, exactly one of A and $X \setminus A$ is in U.

Roughly speaking, one can think of an ultrafilter as equipping X with a notion of "almost all"-ness: relative to U, a subset $A \subseteq X$ contains "almost all" elements of X just in case $A \in U$.

Given an ultrafilter U on some set I (which we will call the set of indices), and an indexed family $\{\mathcal{M}_i : i \in I\}$ of Σ-structures, the ultraproduct $\Pi_U \mathcal{M}_i$ of the family is a Σ-structure defined as follows. First, form the Cartesian product $\Pi_{i \in I} |\mathcal{M}_i|$ of the domains of the structures in the family: that is, each element of $\Pi_{i \in I} |\mathcal{M}_i|$ is an indexed family of elements $\mathbf{a} = \{a^i \in |\mathcal{M}_i| : i \in I\}$, one from each $\mathcal{M}_i$. Then define the following equivalence relation: $\mathbf{a} \sim \mathbf{b}$ iff the set of indices i such that $a^i = b^i$ is in the ultrafilter. That is, $\mathbf{a}$ and $\mathbf{b}$ are equivalent if they are identical at "almost all" models in the family of

---

[16] As discussed above (see footnote 10), the condition of domain-preservation is clearly not very appropriate to the context of scientific theories—but is helpful in simplifying the technicalities.

[17] For an introduction to ultraproducts, see (Hodges, 1997, §8.5).

structures. We then take the domain of the ultraproduct to be the quotient of $\Pi_{i \in I}|\mathcal{M}_i|$ by this equivalence relation (so its elements are sets of "almost-everywhere-identical" indexed families). Finally, we determine the extension of any Σ-symbol R as follows: for n-ary R, the tuple $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$ is in the extension of R if and only if the set of indices i such that $\langle a_1^i, \ldots, a_n^i \rangle \in R^{\mathcal{M}_i}$ is in U. Thus, for instance, an indexed family $\mathbf{a}$ satisfies a unary predicate iff it satisfies the predicate at "almost all" models in the family of structures. An ultraproduct all of whose factors are identical is called an *ultrapower*: an ultrapower of $\mathcal{M}$, relative to the ultrafilter U, will be denoted $\Pi_U \mathcal{M}$.

The key result for ultraproducts is *Łoś's theorem*:[18] given an ultraproduct $\Pi_U \mathcal{M}_i$ and an n-ary Σ-formula ϕ, ϕ is satisfied by the n-tuple $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$ in $\Pi_U \mathcal{M}_i$ if and only if the set of indices i such that ϕ is satisfied by $\langle a_1^i, \ldots, a_n^i \rangle$ in $\mathcal{M}_i$, is in U. Thus, roughly speaking, an ultraproduct satisfies a formula if and only if "almost all" the models in the indexed family satisfy that formula. As a corollary, given some ultrapower $\Pi_U \mathcal{M}$, ϕ is satisfied by the n-tuple $\langle a_1, \ldots, a_n \rangle$ (where every $a_p \in |\mathcal{M}|$) if and only if ϕ is satisfied by $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$ in $\Pi_U \mathcal{M}$, where for every $i \in I$ and $1 \leq p \leq n$, $a_p^i = a_p$.

With this in hand, we can turn to the proof itself.[19]

Proof. The left-to-right direction of the proof is reasonably straightforward: it is just a matter of verifying that for any translation $\tau : T_1 \to T_2$, τ^* preserves isomorphisms, ultraproducts and domains.

Now consider the right-to-left direction. For any model $\mathcal{M}$ of T_2, let $\mathcal{M}^+$ be the definitional expansion of $\mathcal{M}$ to $\Sigma_1 \cup \Sigma_2$ defined by σ: that is, the model such that (i) $|\mathcal{M}^+| = |\mathcal{M}|$; (ii) for any $R \in \Sigma_1$ and any $a_1, \ldots, a_n \in |\mathcal{M}|$, $a_1, \ldots, a_n \in R^{\mathcal{M}^+}$ iff $a_1, \ldots, a_n \in R^{\sigma(\mathcal{M})}$ (exploiting the fact that σ preserves domains); and (iii) for any $S \in \Sigma_2$ and any $b_1, \ldots, b_m \in |\mathcal{M}|$, $b_1, \ldots, b_m \in S^{\mathcal{M}^+}$ iff $b_1, \ldots, b_m \in S^{\mathcal{M}}$. Now define

$$K := \{\mathcal{M}^+ : \mathcal{M} \in \mathrm{Mod}(T_2)\} \tag{7}$$

First, we want to show that K is axiomatisable. A necessary and sufficient condition for this is that K is closed under isomorphism and taking ultraproducts, and has the property that for any structure $\mathcal{A}$, if some ultrapower of $\mathcal{A}$ lies in K then $\mathcal{A}$ itself lies in K.[20] The first two properties follow immediately from σ's preservation of isomorphism and ultraproducts, so it remains only to prove the third.

---

[18]See, for example, (Hodges, 1997, Theorem 8.5.3).

[19]The proof below is a simplified version of the proof (of a more general result) given by van Benthem and Pearce (1984).

[20](Hodges, 1997, Corollary 8.5.13)

To this end, suppose that $\Pi_U \mathcal{A} \in K$. Let us denote the reduct of $\mathcal{A}$ to Σ_2 by $\mathcal{M}$;[21] we want to show that $\mathcal{A} = \mathcal{M}^+$. Taking reducts commutes with taking ultraproducts,[22] so the reduct of $\Pi_U \mathcal{A}$ is $\Pi_U \mathcal{M}$. Since $\Pi_U \mathcal{A} \in K$, $\Pi_U \mathcal{M} \in \mathrm{Mod}(T_2)$. By Łoś's theorem, $\mathcal{M} \in \mathrm{Mod}(T_2)$. So $\mathcal{A}$ is an expansion of a T_2-model. It remains to show that for any $R \in \Sigma_1$ and any $a_1, \ldots, a_n \in |\mathcal{A}|$, $\langle a_1, \ldots, a_n \rangle \in R^{\mathcal{A}}$ iff $\langle a_1, \ldots, a_n \rangle \in R^{\sigma(\mathcal{M})}$. This goes as follows:

$$\begin{aligned}
\langle a_1, \ldots, a_n \rangle \in R^{\mathcal{A}} \quad &\text{iff} \quad \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle \in R^{\Pi_U \mathcal{A}} && \text{(Łoś's theorem)} \\
&\text{iff} \quad \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle \in R^{\sigma(\Pi_U \mathcal{M})} && \text{(since } \Pi_U \mathcal{A} \in K) \\
&\text{iff} \quad \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle \in R^{\Pi_U \sigma(\mathcal{M})} && \text{(}\sigma \text{ preserves ultraproducts)} \\
&\text{iff} \quad \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle \in R^{\sigma(\mathcal{M})} && \text{(Łoś's theorem again)}
\end{aligned}$$

Thus, there is some theory T^+ such that $K = \mathrm{mod}(T^+)$. But now observe that for any two models $\mathcal{A}, \mathcal{B}$ of T^+ (i.e. any $\mathcal{A}, \mathcal{B} \in K$), if the reducts of $\mathcal{A}$ and $\mathcal{B}$ to Σ_2 are identical, then $\mathcal{A} = \mathcal{B}$. Thus, T^+ *implicitly defines* every $R \in \Sigma_1$ in terms of Σ_2. But then by Beth's theorem, T^+ explicitly defines R, and so entails some sentence of the form $\forall x_1 \ldots \forall x_n (R x_1 \ldots x_n \leftrightarrow \tau_R(x_1, \ldots, x_n))$. So, we define an interpretation $\tau : \Sigma_1 \to \Sigma_2$ by setting $\tau(R) = \tau_R$ for every $R \in \Sigma_1$, and extending by requiring τ to commute with the logical constants. It is straightforward to confirm that $\sigma = \tau^*$, and that τ is a translation from T_1 to T_2. $\qquad\square$

Now consider the second question raised at the start of this section: whether the conditions on Theorem 4 can be thought of as imposing a requirement of similarity-preservation. One way to show this would be to argue that an ultraproduct can be regarded as a kind of "limit" of a family of models. In support of this, we could observe that in the context of propositional logic (which, recall, is simply first-order logic where all predicates are nullary) the ultraproduct of a sequence of models is the limit of that sequence in the Stone topology.[23] More generally, Łoś's theorem provides a sense in which an indexed family of models "converges" on the ultraproduct. However, the sense of convergence here must remain rather loose. The standard axioms governing convergence require that if all members of a converging family are identical to some object, then the limit of the family is that object;[24] yet, in general, the ultrapower of a given model is not identical to that model, nor is it even isomorphic to it.[25] Nevertheless,

---

[21] That is, $\mathcal{M}$ is the Σ_2-model such that $|\mathcal{M}| = |\mathcal{A}|$, and for any $R \in \Sigma_2$ and any $b_1, \ldots, b_m \in |\mathcal{M}|$, $b_1, \ldots, b_m \in S^{\mathcal{M}}$ iff $b_1, \ldots, b_m \in S^{\mathcal{A}}$.

[22] (Hodges, 1997, Theorem 8.5.1)

[23] Halvorson and Tsementzis (2017)

[24] See e.g. Dudley (1964), Patten (2014).

[25] Although it is, by Łoś's theorem, elementarily equivalent to it. We could try working with elementary-equivalence classes of models instead, and see if the ultraproduct does provide an appropriate notion of convergence in that setting. However, then there will be the problem of how to justify demanding

I conclude that the above considerations do illuminate some aspects of the relationship between reduction (thought of as an essentially syntactic relation between theories) and supervenience (thought of as an essentially semantic relation between models of theories, or the possible worlds they represent): we have learned that constraints on a supervenience map, stable in purely semantic terms, can enforce reducibility.

References

Barrett, T. W. and Halvorson, H. (2016). Glymour and Quine on Theoretical Equivalence. *Journal of Philosophical Logic*, 45(5):467–483.

Butterfield, J. (2011a). Emergence, Reduction and Supervenience: A Varied Landscape. *Foundations of Physics*, 41(6):920–959.

Butterfield, J. (2011b). Less is Different: Emergence and Reduction Reconciled. *Foundations of Physics*, 41(6):1065–1135.

Dizadji-Bahmani, F., Frigg, R., and Hartmann, S. (2010). Who's Afraid of Nagelian Reduction? *Erkenntnis*, 73(3):393–412.

Dudley, R. M. (1964). On sequential convergence. *Transactions of the American Mathematical Society*, 112(3):483–507.

Fletcher, S. C. (2016). Similarity, Topology, and Physical Significance in Relativity Theory. *The British Journal for the Philosophy of Science*, 67(2):365–389.

Halvorson, H. (2012). What Scientific Theories Could Not Be. *Philosophy of Science*, 79(2):183–206.

Halvorson, H. and Tsementzis, D. (2017). Categories of Scientific Theories. In Landry, E., editor, *Categories for the Working Philosopher*. Oxford University Press, Oxford. References are to draft of July 29, 2015.

Hodges, W. (1997). *A Shorter Model Theory*. Cambridge University Press, Cambridge, UK.

List, C. (2018). Levels: Descriptive, Explanatory, and Ontological. *Noûs*. Forthcoming. References are to the Early View version available at doi: 10.1111/nous.12241.

---

that the supervenience mapping preserve isomorphism.

Nagel, E. (1979). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Hackett, Indianapolis.

Patten, D. R. (2014). *Problems in the Theory of Convergence Spaces*. PhD dissertation, Syracuse University, Syracuse, NY.

Suppes, P. (1957). *Introduction to Logic*. Van Nostrand Reinhold, New York.

Suppes, P. (1967). What is a scientific theory? In Morgenbesser, S., editor, *Philosophy of Science Today*, pages 55–67. Basic Books, New York.

van Benthem, J. and Pearce, D. (1984). A mathematical characterization of interpretation between theories. *Studia Logica*, 43(3):295–303.