Synergic Kinds

Manolo Martínez

Abstract

According to the *homeostatic property cluster* family of accounts, one of the main conditions for groups of properties to count as natural is that these properties be frequently co-instantiated. I argue that this condition is, in fact, not necessary for natural-kindness. Furthermore, even when it is present, the focus on co-occurrence distorts the role natural kinds play in science.

Co-occurrence corresponds to what information theorists call *redundancy*: observing the presence of some of the properties in a frequently co-occurrent cluster makes observations of other members of the cluster comparatively uninformative. Yet, scientific practice often, and increasingly often, singles out as natural groups of properties that are not redundant, but *synergic*: instantiations of properties in synergic clusters are not necessarily informative about instantiations of other properties in the cluster; rather, it is their *joint* instantiation that plays the explanatory role for which the natural kind is recruited.

1 Introduction

Natural kinds are supposed to help us, among other things, explain the frequent success of our inductive practices. In the usual caricature, we are often able to move successfully from *this F is a G* to *all Fs are, likely, Gs*, and at least sometimes these inferences are grounded on the fact that the Fs form a natural kind.[1] The question that will interest me here is how should we construe natural-kindness, so that natural kinds live up to these theoretical expectations.

One popular answer to this question, and perhaps the default answer until the early 90s, is that all natural kinds have an intrinsic *essence*: a set of necessary and jointly sufficient intrinsic properties for an entity to count as a member of the kind in question. The presence of the essence, in its turn, would cause or otherwise explain the presence of all of the properties of Fs we are able to draw successful inductive inferences about. In the usual example of induction, introduced above, it is supposed to work more or less as follows: we observe that a certain F is G. Now, it just so happens that being an F *consists in* having, say, property E—i.e., this property constitutes the essence of F-hood. It also just so happens that the probability of an entity having property G, conditional

---

[1]Not always. See Bird (2015) for discussion.

on its having property E, is very high. Putting all of this together, we are warranted in concluding that Fs are likely to be Gs.[2]

One problem with this appeal to intrinsic essences is that many natural kinds do not have them. The most prominent example is biological species, where the consensus among philosophers of biology is that "the idea that species can somehow be 'defined in terms of their DNA' has no basis in biological fact" (Okasha 2002, p. 197). Yet we are often perfectly justified in inferring, e.g, from the presence of relevant phenotypic traits in a member of a certain species, the likely presence of those same traits in other members of the same species.[3]

The *Homeostatic Property Cluster* [HPC] theory[4], one of the leading contemporary approaches to the metaphysics of natural kinds, is, among other things, an attempt at doing induction without intrinsic essences. According to Boyd's original formulation (most clearly codified in Boyd 1999, p. 143f), HPCs are individuated by clusters (groups) of properties that typically meet a number of conditions, the most important of which are the following two (quoted from Boyd *ibid*):

Clustering as Co-Occurrence: [These properties] are contingently clustered in nature in the sense that they co-occur in an important number of cases.

Homeostasis: Either the presence of some of [these] properties tends … to favor the presence of the others, or there are underlying mechanisms or processes that tend to maintain the presence of [these] properties, or both.

Other conditions make reference to the "causal[5] importance" of this property cluster (i.e., whether effects we care about, theoretically or practically, follow from the co-instantiation of these properties), or the fact that the kind individuated by this cluster has no analytic definition. While these other conditions are important in warranting talk of HPCs as *natural* kinds, the task of explaining how they help ground inductive inference is mainly discharged by the two conditions I have singled out. It should be noted that, in these conditions, no assumptions are made about a privileged essence subset among the properties in the cluster, from which probabilistic connections would flow to all other properties. For all the account says, there just is an unstructured probabilistic connection among properties in the cluster whereby the instantiation of some of them (and none in particular) makes the instantiation of the rest likely.

---

[2]Inductive inference takes widely different forms, of course, and go far beyond this somewhat caricaturesque example. It is possible that some of the shortcomings of the homeostatic property cluster account I identify in this paper could be traced back to relying on this kind of simple induction as a guiding example. As we will see, much of the inductive inference science engages in is significantly more complex.

[3]In current philosophy of biology, essentialism about species has taken a historical turn (Godman 2015; Griffiths 1999; Millikan 1999; Okasha 2002) whereby essences are no longer taken to be intrinsic, but rather historical properties of lineages of populations. In what follows I will be defending a conservative modification of the homeostatic property cluster theory of natural kinds, and in particular one that is compatible with the postulation of historical essences for species.

[4]The original formulation is by Richard Boyd, in his (1989); see also Boyd (1999), Chakravartty (2007), Kornblith (1993), Magnus (2011), among many others.

[5]While throughout this paper I often talk of *causal* structure, the models I will describe in subsequent sections are causally agnostic, and focus on probabilistic (informational) connections. In any case, whenever these connections are grounded on causal facts my discussion applies to them as well. See also footnote 6.

This account is very popular among philosophers working on the metaphysics of natural kinds, but it has certainly not passed unchallenged. One important group of objections comes from a series of papers by Marc Ereshefsky and colleagues (Ereshefsky & Matthen 2005; Ereshefsky 2010; Ereshefsky & Reydon 2015). The main claim developed by Ereshefsky in this body of work is that the HPC account does not agree with actual scientific practice. First, scientific kinds (i.e., those appealed to in the process of scientific inquiry) do not always aim at uncovering causal structure, while the avowed focus of the HPC account is precisely on "the accommodation of inferential practices to relevant causal structures" (Boyd 1999, p. 159; cited in Ereshefsky & Reydon 2015, p. 973). Second, in the HPC account, the entry point to the process of delineating causal structure is *similarity* (i.e., property overlap) among kind instances: homeostatic clusters go hand in hand with sets of kind instances in which the same collection of properties is (noisily) replicated. Yet for many natural kinds similarity among instances is, at best, of secondary importance: for example, biological species, precisely one of the chief motivating cases for the HPC account, are often polymorphic (say, sexually dimorphic, as with mammals), and the similarity among morphs need not be particularly high, and on occasion is extremely low (Magnus 2011; Martínez 2015). In such *heterostatic kinds*, (kinds with stably dissimilar instances; Ereshefsky & Matthen 2005) natural-kindness and similarity come apart.

In this paper I use simple probabilistic models to examine whether the *Clustering as Co-Occurrence* and *Homeostasis* conditions at the heart of the HPC account are able to ground inductive inference, whenever natural kinds are called upon to do so. The upshot of my discussion will be that, even granting Ereshefsky's first main point that many *bona fide* scientific kinds are not covered by the HPC account, and even if we focus on the delineation of causal structure, the account does not deliver: a great deal of causal structure (and, correspondingly, very many inferential processes) are not captured by homeostatic property clusters, as traditionally conceived. On the other hand, I will characterize a different ingredient in natural-kindness, alongside co-occurrence, able to accommodate the bits of structure that traditional HPC accounts leave out. In particular, I will show that heterostatic kinds, *pace* Ereshefsky, can ground inductive inference, and are compatible with a view of natural kinds as entities that delineate causal structure.

More concretely, I will argue that meeting *Clustering as Co-Occurrence* is, in fact, not necessary for a property cluster to ground inductive inference. I will also argue that, even when this condition is met, focusing on co-occurrence distorts the role natural kinds often play in inductive inference. Co-occurrence corresponds to what information theorists call *redundancy*: properties in a cluster are redundant insofar as observing the presence of some of them makes observations of the rest comparatively uninformative; but scientific practice often (and increasingly often) singles out as natural groups of properties that are not (or not just) redundant, but *synergic*. Instantiations of properties in synergic clusters are not necessarily informative about instantiations of other properties in the cluster; rather, it is the joint instantiation of all or many of those properties that plays the explanatory role for which the natural kind is recruited.

In sections 2 and 3 I present a series of probabilistic models where accounts of natural-kindness can be given formal expression. I will use them to construct idealized examples

of essence kinds, HPCs, and synergic, largely non-redundant (yet natural) kinds. These latter synergic kinds are offered as counterexamples to the implicit HPC claim that homeostatic co-occurrence is enough to accommodate all inference-relevant causal structure. In section 4 I examine a rejoinder to this counterexample, inspired by Boyd's treatment of species polymorphism, that I interpret as aiming at redescribing synergy in terms of redundancy. I conclude that this redescription relies on metaphysically suspect claims, and is anyway theoretically idle.

While the bulk of the discussion in this paper is framed in terms of idealized models, section 5 shows that synergic kinds are not just an academic exercise: I briefly describe two contemporary research programs, in genetics and neuroscience, that are predicated on the existence of natural kinds with an important synergic component. Finally, section 6 offers some concluding remarks, and a tentative substitute for the *Clustering as Co-Occurrence* and *Homeostasis* conditions. The resulting theory is a conservative extension of traditional HPC accounts.

2   Probabilistic Models of Natural-Kindness

As the discussion above suggests, much of the contemporary debate on the metaphysics of natural kinds (and, at any rate, the aspect of the debate to which this paper aims at contributing) focuses on which are the criteria for counting groups of properties, instantiations of which bear different probabilistic relations to one another, as natural. I will investigate this question through probabilistic models.

To keep it as simple as possible, I will focus on "worlds" in which only six properties could possibly be instantiated. Each of these properties will be modeled by a binary random variable, $P_1 \ldots P_6$, that takes the value 0 if the corresponding property is not instantiated, and 1 if it is. For simplicity, I will sometimes abuse language and use P_i to refer to the properties themselves, not the associated random variables.

Events in these worlds are represented by the joint values of the six random variables. For example, 000111 stands for the event consisting of properties four to six, and no other, being instantiated. Now, facts concerning which properties are likely, or not, to co-occur, which properties are likely instantiated conditional on others, etc. can be summarized in a joint probability distribution over those six random variables. Different worlds will correspond to different distributions. Our task is to work out which groups of properties should count as natural in these worlds. To get the hang of things, let's see how this works for essence kinds.

2.1   Essence Kinds

Take, to start with, World 1 as described by the probability distribution in Figure 1. In the figure, the probability distribution is summarized by a Bayesian network: a graph that is both acyclic (i.e., with no loops), and directed (i.e, where edges have a privileged

direction). In Bayesian networks, if an edge goes from node A to node B we say that A is a parent to B (and B a child to A). The graph represents the following so-called "Markov condition": every node n_i in the network is independent of any other node n_j, conditional on the value of n_i's parents. This means that, given the Bayesian network, and, for every node, a table of probabilities of that node conditional on its parents, we can reconstruct the full joint probability distribution. For much more on Bayesian networks, see Koller & Friedman (2009), chapter 3.

In this world there are two "root" properties, P_1 and P_4, instantiations of each of which reliably (if noisily) correlate with the instantiations of two other "children" properties— P_2 and P_3 for P_1; P_5 and P_6 for P_4. P_1 and P_4 are independent from one another.

a	$P(P_5 = 1 \vert P_4 = a)$	$P(P_5 = 0 \vert P_4 = a)$
0	0.05	0.95
1	0.95	0.05

$P(P_4 = 0)$	$P(P_4) = 1$
0.5	0.5

a	$P(P_6 = 1 \vert P_4 = a)$	$P(P_6 = 0 \vert P_4 = a)$
0	0.05	0.95
1	0.95	0.05

a	$P(P_2 = 1 \vert P_1 = a)$	$P(P_2 = 0 \vert P_1 = a)$
0	0.05	0.95
1	0.95	0.05

$P(P_1 = 0)$	$P(P_1 = 1)$
0.5	0.5

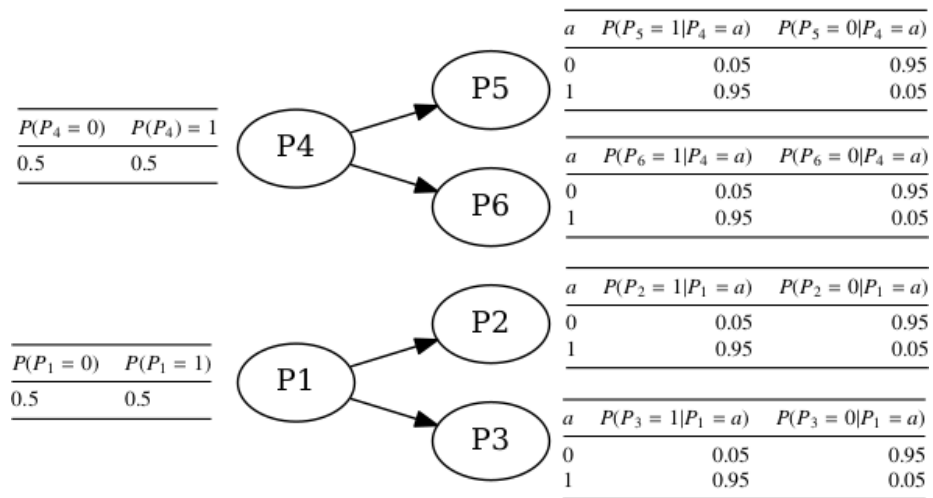| a	$P(P_3 = 1 \vert P_1 = a)$	$P(P_3 = 0 \vert P_1 = a)$
0	0.05	0.95
1	0.95	0.05

Figure 1: Two essence kinds

Faced with this world, we want to know which groups of properties should be considered natural. That there *should* be natural groups of properties seems clear: this world is able to accommodate induction within certain groups of properties (for example, witnessing an instantiation of P_2 provides very good evidence that properties P_1 and P_3 will be instantiated as well; the same happens with properties in the P_4, P_5, P_6 group) but not across groups (nothing of what we learn about, e.g., P_2 gives us any information about, e.g., P_5.)

In this particular case, an essentialist account of natural-kindness is compelling. There is an asymmetry built into the probability distribution, whereby probability of instantiation "flows" from root properties onto their children, in the following sense: first, once we know, e.g., the value of the P_1 random variable, P_2 is no longer informative about P_3, and vice versa (this is just the Markov condition on Bayesian networks—see above.) Second, if we have to choose the instantiation of one property among P_1, P_2 and P_3 as the criterion for the instantiation of the other two, it is P_1 that minimizes false alarms and misses in the instantiations of the other properties (e.g., $P(P_2 \vert P_1) = 0.95$, but

$P(P_2|P_3) = 0.905$.) *Mutatis mutandis* for P_4, P_5 and P_6. It makes sense, then, to see World 1 as composed of two independent essence kinds, one with P_1 as essence, another with P_4 as essence.[6]

So far, so good, but, as we are about to see, there are other probabilistic structures in which there are equally compelling candidates for the role of natural kinds, but in which no one property, or group thereof, has a claim to being the essence of the kind. The homeostatic property cluster account was formulated with these other structures in mind. The following subsection provides an example.

2.2 Redundant Kinds

Consider now World 2, described by the probability distribution in Figure 2, together with the factors in Table 1. In Figure 2, the probability distribution is summarized by a *Markov network*, In these graphs, probabilistic influence diffuses in no privileged direction, and the full joint distribution is constructed by multiplying the so-called "factors" that record pairwise influences between connected nodes, then normalizing. For much more on Markov networks, see Koller & Friedman (2009), chapter 4.

This is a world with two clusters, $C_1 = \{P_1, P_2, P_3\}$ and $C_2 = \{P_4, P_5, P_6\}$, such that properties within each one of them co-occur frequently, but instantiations of properties in a cluster are independent from those of properties in the other cluster.

P_i	P_j	
0	0	99
1	0	1
0	1	1
1	1	99

P_i	P_j	
0	0	25
1	0	25
0	1	25
1	1	25

Table 1: The factor in the left table records the way in which properties P_i and P_j within the same cluster influence one another: co-occurrence (and co-absence) are vastly more probable than situations in which one of the two properties is instantiated but not the other. The factor in the right table records the way in which properties P_i and P_j influence one another when they belong to different clusters (i.e., $P_i \in C_1$ and $P_j \in C_2$): all rows are equal, that is, trans-cluster properties are independent from one another.

Here we have as much potential for induction as we had in World 1: instantiations of properties in C_1 (C_2) go hand in hand with one another and, for example, seeing an instantiation of P_5 provides excellent evidence that P_4 and P_6 will be instantiated too. Yet, this world provides no grounds for the postulation of essences: there is no

---

[6]Note that I have not given a causal gloss on the role of essences. This is because causal facts cannot be read off probability distributions. This causal agnosticism is, I submit, an advantage of the model: it allows it to apply to the HPC account proper, with its reliance on causal connections, and to other, related accounts such as Slater (2015)'s *stable property clusters*, that substitute causal connections with a modal stability constraint. This kind of constraints, arguably, are precisely what probability distributions inform us of. I discuss Slater's account in the following subsection.
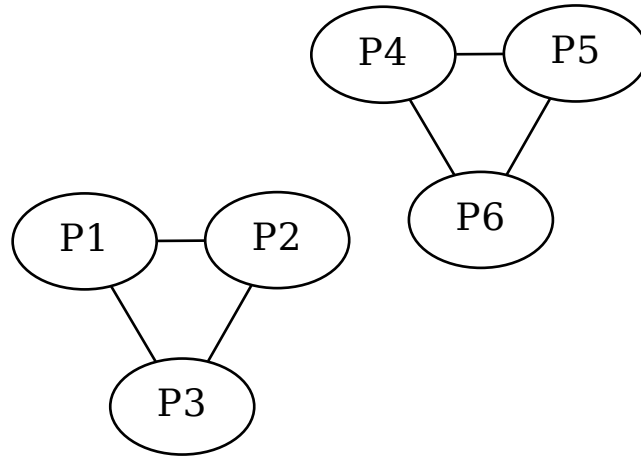
Figure 2: Two HPCs

asymmetry in the pattern of instantiations of properties within each cluster that might justify such postulation (and the probability distribution itself is causally agnostic).

The HPC suggestion is that we abandon the search for privileged properties, and simply focus on which groups of properties frequently co-occur with one another. The partition of World 1 in clusters C_1 and C_2 is, plausibly, then, the most natural one—the one cutting at the proverbial joint. Within each cluster, properties are frequently co-instantiated with one another, but not with those in the other cluster. This is mirrored in the fact that induction works within each of these two clusters, but not across them. No other partition of World 1 into clusters has this property.

This probabilistic relation among properties belonging to the same natural kind (properties within C_1 or C_2, in our example) is what the HPC account, and other related ones, single out as one of the main hallmarks of natural-kindness, under the name of *co-occurrence*. It is also a sufficient condition for what information theorists call *redundancy* (Griffith et al. 2014; Williams & Beer 2010): properties in a frequently co-occurring cluster tend to be doing the same thing at the same time; this means that a big part of the information that individual properties carry about World 2 is "repeated" within a cluster; one can learn from an instance of P_5 that all of C_2 is likely to be instantiated, and subsequent exposure to instances of P_4 and P_6 is, to that extent, comparatively informationally-redundant: roughly, the information that witnessing the instantiation of several properties gives about the world is less than the sum of the information that each

such instantiation gives on its own.[7]

Matthew Slater's *stable property cluster* account of natural kinds (2015), has a more formal bent and its reliance on informational redundancy is, thus, particularly clear. According to Slater, the main condition to be met by natural clusters of properties is what he calls *cliquish stability*:

> ... properties are clustered in such a way that possession of some of them reliably (if imperfectly) indicates the possession of the whole cluster (Slater 2015, p. 397)

Groups of properties in such clusters are reliable indicators of the whole. Like a real-life clique, such clusters are *boring*:

> Peg, Quinn, Ralph, Sarah, and Tim form a clique, say. Spotting Peg, Quinn, and Ralph at the mall means that Sarah and Tim are probably there as well. (Slater 2015, p. 397)

Peg's gang are predictable, and this makes them a suitable target for induction. This is the reason why redundant clusters figure so prominently in the HPC story, but this fixation on redundancy stems from too narrow a focus on the (undoubtedly central) case in which the causal structure on which induction depends is of the sort that gives rise to sets of similar things—things which, more or less noisily, reinstantiate the very same set of properties, as in worlds 1 and 2.

As I said in the introduction, Ereshefsky & Reydon (2015), among others, have criticized the almost exclusive focus on similarity in the HPC account: scientific taxonomies often prioritize other criteria. Ereshefsky's discussion, though, suggests that this is because scientific taxonomy has other objectives than just uncovering the grounds of our inferential practices. This tacit equation between similarity- (and thus redundancy-) based taxonomies and the explanation of inference is unhelpful. Inference in science can, and routinely does, go well beyond what similarity affords. Even if we are solely interested in the accommodation of inferencial practices, then, our account of natural-kindness needs to go beyond redundancy (co-occurrence, cliquish stability, etc.) In the following section, World 3 drives this point home.

3   Synergic Kinds

Consider now World 3, as described by the Bayesian network in Figure 3.

World 3 consists in two independent, noisy, exclusive-OR [XOR] logical gates: that is, P_5 is instantiated when one but not both of P_4 and P_6 is instantiated (plus some noise). *Mutatis mutandis* for P_2, P_1 and P_3. This world is not at all about redundancy

---

[7]More formally, two properties P_i and P_j provide redundant information about the world iff the mutual information between the joint random variable, W, of all random variables in the world, and the joint random variable (P_i, P_j) is less than the sum of the mutual informations of W and each individual random variable: $I(W; P_1) + I(W; P_2) > I(W; P_1, P_2)$. The definition of redundancy for more than two random variables is an open theoretical problem.
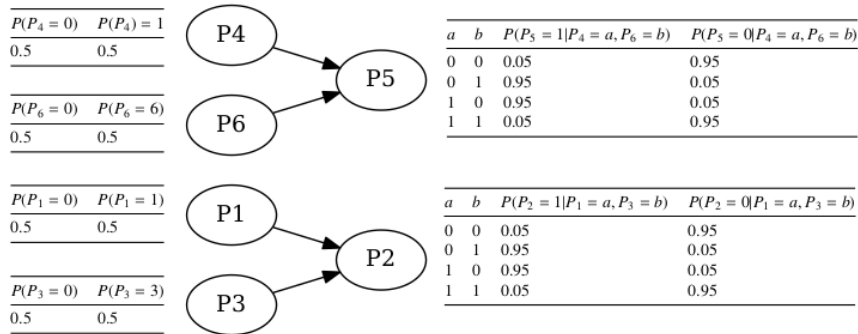
$P(P_4 = 0)$	$P(P_4) = 1$
0.5	0.5

$P(P_6 = 0)$	$P(P_6 = 6)$
0.5	0.5

P4 → P5 ← P6

a	b	$P(P_5 = 1 \mid P_4 = a, P_6 = b)$	$P(P_5 = 0 \mid P_4 = a, P_6 = b)$
0	0	0.05	0.95
0	1	0.95	0.05
1	0	0.95	0.05
1	1	0.05	0.95

$P(P_1 = 0)$	$P(P_1 = 1)$
0.5	0.5

$P(P_3 = 0)$	$P(P_3 = 3)$
0.5	0.5

P1 → P2 ← P3

a	b	$P(P_2 = 1 \mid P_1 = a, P_3 = b)$	$P(P_2 = 0 \mid P_1 = a, P_3 = b)$
0	0	0.05	0.95
0	1	0.95	0.05
1	0	0.95	0.05
1	1	0.05	0.95

Figure 3: Two synergic kinds

(co-occurrence, cliquish stability, etc.): no property reliably indicates the presence of any other property or subcluster. Take, for example, P_1: knowing that this property is instantiated gives us *no* information about the state of any other property in the world. The same happens with any other single property. That is, for all $i, j \neq i$, $P(P_i = 1 \mid P_j = 1) = P(P_i = 1)$.

Still, this world is perfectly able to accommodate inductive inference. Within each of the two clusters $C_1 = \{P_1, P_2, P_3\}$ and $C_2 = \{P_4, P_5, P_6\}$ seeing instances of two of the properties in a cluster provides very good inductive evidence that the third property of the cluster is *not* instantiated. Alternative partitions of the world into clusters (say, even-numbered and odd-numbered) fail to accommodate these, or any other, inferences. This suggests that C_1 and C_2 are just as natural in World 3 as they were in Worlds 1 and 2, and that World 3 has as conspicuous a joint as the other ones.

The main difference between the clusters in Worlds 1 and 2, and the clusters in World 3 is that what makes the latter be natural units is mostly *synergic*, as opposed to redundant, information (Anastassiou 2007; Bertschinger et al. 2013): each property individually carries no information about the others on its own, but the joint instantiation of each two of them does. This is synergy, roughly, because the information about the world obtained by witnessing the instantiation of several properties is higher than the sum of the information provided by each separate instantiation.[8] Even then, they do not indicate that the third property in the cluster will likely be instantiated too, quite the contrary: the case in which all properties in the cluater are instantiated (the typical member of an HPC kind) has very low probability. I will talk of *synergic kinds* to refer to natural

---

[8]More formally, two properties P_i and P_j provide synergic information about the world iff the mutual information between the joint random variable, W, of all random variables in the world, and the joint random variable (P_i, P_j) is more than the sum of the mutual informations of W and each individual random variable: $I(W; P_1) + I(W; P_2) < I(W; P_1, P_2)$. The definition of synergy for more than two random variables is also an open theoretical problem.

kinds, such as the ones in World 3, that have a significant synergic component.

Slater invites us to picture a clique, a group of friends that hang out together most of the time, as an intuitive example of a stable cluster kind. A love triangle offers an analogously intuitive example for synergic kinds: Ralph and Peg are best friends with one another, and both are in love with Quinn. As a result, the three of them hardly ever hang out (that would be awkward) but any two of them very often do. If you know that they form a love triangle, you can leverage this information: if you are looking for Peg and you see Ralph and Quinn at the mall, you have excellent inductive evidence that she will *not* be around, and you should go looking elsewhere.

4   Redundant redescriptions

Synergic kinds are as capable of grounding inductive inference as purely redundant ones. If, moreover, they meet the other conditions singled out by HPC theorists (causal importance, existence of a natural kind term, etc. see Boyd (1999), and above) then they are natural kinds if traditional HPCs are. The exclusive focus on redundancy in HPC theorizing stems, I have hypothesized, from too narrow a view of the forms inductive inference can take. Richard Boyd's discussion of species polymorphism is a good example of this. Many, perhaps most, biological species present a discrete set of stably different phenotypic variants, or "morphs". Mammals, for example, are sexually dimorphic: they come in two different phenotypic variants, male and female. As Ereshefsky & Matthen (2005) helpfully put it, such species are *heterostatic*: it is not that members of a species *fail* to be perfectly similar and, e.g., some lions happen to have a mane and some happen not to; rather, it is a stable characteristic of certain species that they come into a small number of sets of things that are comparatively similar to those in the same set, and comparatively dissimilar to those in other sets.

I will presently suggest that, at least sometimes and at least partly, the heterostasis in polymorphic species needs to be captured in terms of synergic property interactions in the species in question. Here, however, is Boyd making do with just redundancy:

> The fact that there is substantial sexual dimorphism in many species and the fact that there are often profound differences between the phenotypic properties of members of the same species at different stages of their life histories ... together require that we characterize the homeostatic property cluster associated with a biological species as containing lots of conditionally specified dispositional properties for which canonical descriptions might be something like, "if male and in the first molt, P," or "if female and in the aquatic stage, Q." (Boyd 1999, p. 165)

Boyd's point, in essence, is that we can increase the redundancy of an arbitrary probability distribution by adding random variables. To see how this works consider again the Peg-Quinn-Ralph love triangle introduced above. One way to see it is as a polymorphic kind: whenever you encounter the love triange, it's under the form of one out of

three stably different "morphs"; Peg-Quinn, Peg-Ralph, and Quinn-Ralph. Boyd's suggested treatment of polymorphic kinds amounts to introducing three new "conditionally specified dispositional properties" for the love triangle to have:

- CS_1: the property of being such that, if Peg is there and Quinn is there, then Ralph is not there.
- CS_2: the property of being such that, if Peg is there and Ralph is there, then Quinn is not there.
- CS_3: the property of being such that, if Quinn is there and Ralph is there, then Peg is not there.

and then pointing out that all three morphs have all three new properties. For example, when it is Ralph and Quinn that are together at the mall, the love triangle still has the property that if Peg is (were) there and Quinn is there, then Ralph is not (would not be) there, etc. Once augmented with these three properties, the combinations of properties of the triangle that concentrate most of the probability is given by Table 2.

Table 2: The "augmented" love triangle. P, Q and R are random variables associated with Peg, Quinn and Ralph being at the mall, respectively. The CS_i are the Boydian conditionally specified dispositional properties introduced above.

P	Q	R	CS_1	CS_2	CS_2
1	1	0	1	1	1
1	0	1	1	1	1
0	1	1	1	1	1

These properties have made the love triangle largely informationally redundant: five out of six properties are robustly tokened together when a "member" of the kind is present, and this is probably enough to move it into HPC territory. The thing is, the three rightmost columns in Table 2 are doing no real theoretical work; they just shadow the leftmost three, and, in fact, if we were to infer the absence of Ralph from the presence of Peg and Quinn, by using the new properties we would have to "read into" CS_1—that is, we would need to consult the other three columns to find out how R behaves when P and Q are both instantiated. That is to say, while it is certainly true that P, Q and R are related in the way that the three CS_i properties say they are related, this is *a property of the whole kind, not of each individual morph*.

Back to species polymorphism proper, the Boydian postulation of conditionally specified properties amounts to claiming that, e.g., triplewart seadevils (*Cryptopsaras couesii*) are extremely sexually dimorphic in virtue of putative facts such as this: each individual male seadevil (Ken, say), *had he been a female*, would have had a fully formed digestive system. Whether Ken himself could have been a female triplewart seadevil is a question best left to metaphysicians,[9] and one that has no bearing on the status of sexual dimorphism (cf. Magnus 2011).

---

[9]The answer to this question might have to do, for example, with whether Ken has an individual essence,

Recognizing kinds with synergic components alongside purely redundant ones allows us to see dimorphism as a structural property of the *probability distribution* associated with such species. With quite a bit of idealization, this would work as follows. We can recognize three subclusters in the triplewart seadevil property cluster:[10] A *COMMON* subcluster, composed of properties (say, genetic, epigenetic or behavioral) typically common to both morphs; a *MALE* subcluster, composed of properties typically exclusive of the male morph; and a *FEMALE* subcluster, composed of properties typically exclusive of the female morph. These three subclusters are related in a synergistic manner, reminiscent of love triangles: if a morph has properties in *COMMON* and in *FEMALE*, then it will (typically) not have properties in *MALE* (female seadevils do not have properties that are exclusive of the male morph); *mutatis mutandis* for morphs with properties in *COMMON* and *FEMALE*. Finally, if a morph has many properties in both *MALE* and *FEMALE* then it typically will not be a seadevil, and thus will not have the properties in *COMMON*.

Although a more realistic treatment of the seadevil property cluster would surely uncover a much more complex probabilistic structure, hopefully the foregoing sketch is suggestive of the way in which synergy can help describe polymorphic, heterostatic kinds. Polymorphism lies in the way morphs are informationally related to one another, not in the exotic metaphysical potentialities of individual morphs (see Martínez 2015 for a fuller account of species polymorphism along these lines.)

5   Synergic Kinds in Science

So far I have argued against the necessity of homeostatic co-occurrence for natural-kindness using mostly formal models. Synergic kinds, on the other hand, are far from a mere academic curiosity. Before wrapping up, I will quickly point to two research programs in which kinds with synergic components are postulated rather explicitly.

The first example is research on *epistasis* in molecular genetics (Cordell 2002; Fish, Capra & Bush 2015; Mackay 2014; Moore 2003; Watkinson et al. 2008). Epistasis consists in "non-linear molecular interactions [underpinning] the genotype-phenotype map" (Mackay 2014, p. 22). While epistatic effects have been known since the beginning of the twentieth century (Mackay, Stone & Ayroles 2009, p. 565), it's only recently that their systematic study has become computationally feasible.

In epistasis the contribution of different pieces of genetic material to the expression of phenotypic traits is often synergic, in the sense explained above: the analogues of the properties in the model worlds discussed in previous sections would be the presence or absence of certain pieces of genetic material, together with the presence or absence of a

and whether such essence is independent of his sex. Adherents to Kripkean *essentiality of origin* theses (Kripke 1980) will answer this question affirmatively if triplewart seadevil sex is fixed after conception, but there might be other reasonable positions. The worry I am sketching here, and is developed more fully by Magnus in his (2011), is that Boyd's approach makes polymorphism hostage to these comparatively arcane metaphsical considerations.

[10]This discussion is indebted to an anonymous referee.

certain phenotypic complex of interest (say, the one associated with a certain disease.) The non-linearity associated with epistasis will often mean that the information that the joint instantiation of different pieces of genetic material carry about the presence or absence of the phenotype will be different from the sum of the individual informational contributions of these pieces. One immediate consequence of this is that those traits for which epistatic interaction is important will come out synergic. Take, for example, fruit-fly wings: plausibly, they form a natural kind, yet we know that their shape is underwritten by epistatic effects (Weber et al. 1999, 2001). This means that FRUIT-FLY WING will be a synergic kind, and not a traditional HPC.

The wings of fruit flies provide an example among many of a kind, antecedently recognized as natural, that happens to have a synergic component, as a result of the way in which genotype carries information about phenotype. As a different sort of example,[11] research on epistasis is prompting geneticists to characterize *new* putative natural kinds, almost entirely based on synergic connections. E.g., it is suggested that "gene interaction networks" (also GINs henceforth) in model organisms such as fruit flies or *Sacharomyces cerevisiae*, are "likely to be generalizable ... in other species" (Mackay 2014, Box 1). One natural way of reading this suggestion is as claiming that GENE INTERACTION NETWORK is likely to be a natural kind. The properties of GINs that are postulated to be generalizable across species include, among others, their being small-world networks[12], or that fitness of double mutants depend on mutant genes being in the same or different pathways (Mackay, *ibid.*). These properties go well beyond what a co-occurrence based architecture can describe: if a GIN is, in particular, a small-world network, then it is not just a redundant cluster. If one wishes to infer phenotypic traits from genotypic data, playing fast and loose with network architecture just won't do.

A possible rejoinder to this second sort of example on behalf of the HPC theorist is that what Mackay is suggesting in the *op. cit.* is, precisely, that GINs are HPCs, with properties such as *being a small-world network* being part of the property cluster in question. But, as I argued in section 4 in the context of species polymorphism, if we wish to use our knowledge of the GENE INTERACTION NETWORK kind to infer the behavior of a new, unknown GIN, we will have to "read into" such distribution-summarizing properties. The claim that most GINs are small-world (just as the claim that all mammals are sexually dimorphic) is most useful when understood as a constraint on the probability distribution of any new GIN (mammal species), not as a property to add to a putative cluster of co-occurrent properties.

The second example of a research program in which synergy is routinely appealed to is brain *connectomics* (Alivisatos et al. 2012; Hagmann et al. 2010; Sporns, Tononi & Kötter 2005; Van Dijk et al. 2010), which is aimed at describing the connectivity network (the *connectome*) of different nervous systems, under the hypothesis that "the pattern of elements and connections as captured in the connectome places specific constraints on brain dynamics, and thus shapes the operations and processes of human cognition" (Sporns, Tononi & Kötter 2005, p. 249). What we know about the human connectome

---

[11]I would like to thank an anonymous referee for helping me distinguish these two sorts of examples.

[12]That is, graphs in which individual nodes have a small average number of connections to other nodes, yet the average distance between two arbitrary nodes in the graph is also small. See Watts & Strogatz (1998).

suggests, unsurprisingly, that "the human brain is a highly complex organ with a great number of structurally distinct, heterogeneous, yet interconnected components" (*op. cit.*, p. 246). Beyond the coarsest level of description, then, a characterization of the HUMAN BRAIN kind will be more complicated, and have considerably more structure, than the mere agreggation of co-occurrent properties in a cluster. HUMAN BRAIN, I suggest, is another example of a natural kind (antecedently recognized as such) that happens to be highly informationally synergic. Suppose, for the sake of the exercise, that the relevant properties in the human brain that should figure in a probabilistic model of the sort developed in sections 2 and 3 are related to the presence or absence of particular cortical minicolumns[13]. The claim then is that the activities related to such properties depend on one another in a highly synergic fashion, and that mere co-occurrence of minicolumn-related activity will not be able to capture these dependences.

Martjn van der Heuvel and colleagues' *comparative connectomics* aims at uncovering "conserved themes of wiring" in cross-species brain network topology (Heuvel, Bullmore & Sporns 2016, p. 345). Here, as it happens in research on epistasis, and for the same reasons, the restrictions that these wiring themes place on connectomes are not capturable without residue as redundancy. Here, too, research in connectomics appears to be postulating the existence of a CONNECTOME kind, analogous to the putative GENE INTERACTION NETWORK kind discussed above. In this connection, it bears mentioning that a very prominent psychological theory of the nature of consciousness (Tononi and colleagues's *Integrated Information Theory* of consciousness; Edelman & Tononi 2013; Tononi & Edelman 1998; Tononi 2004) explicitly identifies consciousness (perhaps a natural kind in its own right) with highly synergic subsets of the human connectome.

Finally, it should be noted that in the field of behavioral genetics both of the above research programs, with their attendant synergic kinds, come together. Schaffner's excellent (2016) book gives a good sense of the complexities that result from this interaction. Putative natural kinds in behavioral genetics (SCHIZOPHRENIA, perhaps) are unlikely to be classical HPCs.

6   Conclusion

The homeostatic property cluster account provides a good approximation to the way in which natural kinds ground inductive inference in many central cases. On the other hand, I have argued, it is incomplete: the sorts of probabilistic relations among properties that sustain inductive inference go well beyond what co-occurrence affords. At least in some situations, these other relations among properties will warrant the description of kinds which are as natural as HPCs.

I have suggested that connectomics, and research on epistatic effects on complex traits, offer examples of what I have called synergic kinds. These two research programs have taken off only when the analysis of massive datasets has started to be computationally

---

[13]Suppose, that is, that cortical minicolumns are the atomic functional units in a mammalian cerebral cortex. See Sporns, Tononi & Kötter (2005), p. 247, for discussion.

feasible. It is perhaps significant that Richard Boyd's original formulation of the HPC account, in the early 90s, predates these developments. In any event, they show the need for a reformulation of the HPC account.

I have argued that this reformulation should recognize a synergic component in the informational glue that binds properties together in a natural-kind cluster—over and above the redundant component on which HPC theorists have single-mindedly focused. That the mutual information between random variables can be, to varying degrees, redundant and synergic is, of course, not new to information theorists,[14] but the significance of this fact for HPC-style accounts of natural-kindness has so far been overlooked.

What, then, should substitute co-occurrence in a broadly HPC-style account of natural-kindness? One natural suggestion is the smallest genus of which redundancy and synergy are both species: informational *dependence*, quite simply. The following condition aims at capturing this, and is offered as a replacement for the *Clustering as Co-Occurrence* and *Homeostasis* conditions in Boyd's original formulation. A set of properties, *F*, will count as natural if, among other things,

Informational Dependence: Properties in *F* are *informationally connected* in nature in the sense that, in an important number of cases, F cannot be partitioned into two informationally independent subsets.[15]

Some of the other conditions that a natural group of properties has to meet have been spelled out in the papers cited in the introduction, and many others. My *Informational Dependence* condition is designed to be the smallest possible departure from HPC orthodoxy that respects the facts about scientific inquiry discussed in this paper, and I expect that it can be readily plugged in the larger HPC picture. The main change urged in this paper is sensitivity to the fact that, when two random variables are not perfectly independent, their connection can be leveraged for inferential purposes. Any kind of glue, not just co-occurrence, will do.

Acknowledgements

---

[14]Although the precise way in which information should be separated into its synergic and redundant components *is* an open theoretical problem, and one which is the focus of much recent research. See Williams & Beer (2010); Griffith et al. (2014); Bertschinger et al. (2013); Griffith & Koch (2014).

[15]No analogue of the *Homeostasis* condition is needed: the presence of informational dependence already presupposes that probabilistic connections are not a matter of chancy coinstantiations.

References

Alivisatos, AP, Chun, M, Church, GM, Greenspan, RJ, Roukes, ML & Yuste, R 2012, 'The brain activity map project and the challenge of functional connectomics', *Neuron*, vol. 74, no. 6, pp. 970–974.

Anastassiou, D 2007, 'Computational analysis of the synergy among multiple interacting genes', *Molecular systems biology*, vol. 3, no. 1, p. 83.

Bertschinger, N, Rauh, J, Olbrich, E & Jost, J 2013, 'Shared information - new insights and problems in decomposing information in complex systems', *Proceedings of the European Conference on Complex Systems 2012*, Springer, pp. 251–269.

Bird, A 2015, 'The metaphysics of natural kinds', *Synthese*, pp. 1–30.

Boyd, R 1999, 'Homeostasis, Species, and Higher Taxa', in RA Wilson (ed.), *Species: New Interdisciplinary Essays*, Mit Press, pp. 141–185.

Boyd, R 1989, 'What Realism Implies and What It Does Not', *Dialectica*, vol. 43, no. 1-2, pp. 5–29.

Chakravartty, A 2007, *A Metaphysics for Scientific Realism: Knowing the Unobservable*, Cambridge University Press.

Cordell, HJ 2002, 'Epistasis: What it means, what it doesn't mean, and statistical methods to detect it in humans', *Human molecular genetics*, vol. 11, no. 20, pp. 2463–2468.

Edelman, GM & Tononi, G 2013, *Consciousness: How matter becomes imagination*, Penguin UK.

Ereshefsky, M 2010, 'What's wrong with the new biological essentialism', *Philosophy of Science*, vol. 77, no. 5, pp. 674–685.

Ereshefsky, M & Matthen, M 2005, 'Taxonomy, Polymorphism, and History: An Introduction to Population Structure Theory', *Philosophy of Science*, vol. 72, pp. 1–21.

Ereshefsky, M & Reydon, TAC 2015, 'Scientific kinds', *Philosophical Studies*, vol. 172, no. 4, pp. 969–986.

Fish, A, Capra, JA & Bush, WS 2015, 'Are Genetic Interactions Influencing Gene Expression Evidence for Biological Epistasis or Statistical Artifacts?', *bioRxiv*, p. 020479.

Godman, M 2015, 'The Special Science Dilemma and How Culture Solves It', *Australasian Journal of Philosophy*, vol. 93, no. 3, pp. 491–508.

Griffith, V & Koch, C 2014, 'Quantifying synergistic mutual information', *Guided Self-Organization: Inception*, Springer, pp. 159–190.

Griffith, V, Chong, EK, James, RG, Ellison, CJ & Crutchfield, JP 2014, 'Intersection

information based on common randomness', *Entropy*, vol. 16, no. 4, pp. 1985–2000.

Griffiths, PE 1999, 'Squaring the Circle: Natural Kinds with Historical Essences', in RA Wilson (ed.), *Species: New Interdisciplinary Essays*, Mit Press, pp. 209–28.

Hagmann, P, Cammoun, L, Gigandet, X, Gerhard, S, Grant, PE, Wedeen, V, Meuli, R, Thiran, J-P, Honey, CJ & Sporns, O 2010, 'MR connectomics: Principles and challenges', *Journal of neuroscience methods*, vol. 194, no. 1, pp. 34–45.

Heuvel, MP van den, Bullmore, ET & Sporns, O 2016, 'Comparative Connectomics', *Trends in Cognitive Sciences*, vol. 20, no. 5, pp. 345–361.

Koller, D & Friedman, N 2009, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press.

Kornblith, H 1993, *Inductive Inference and its Natural Ground: An Essay in Naturalistic Epistemology*, The MIT Press.

Kripke, S 1980, *Naming and Necessity*, Blackwell.

Mackay, TF 2014, 'Epistasis and quantitative traits: Using model organisms to study gene-gene interactions', *Nature Reviews Genetics*, vol. 15, no. 1, pp. 22–33.

Mackay, TF, Stone, EA & Ayroles, JF 2009, 'The genetics of quantitative traits: Challenges and prospects', *Nature Reviews Genetics*, vol. 10, no. 8, pp. 565–577.

Magnus, PD 2011, 'Drakes, Seadevils, and Similarity Fetishism', *Biology and Philosophy*, vol. 26, no. 6, pp. 857–870.

Martínez, M 2015, 'Informationally-Connected Property Clusters, and Polymorphism', *Biology and Philosophy*, vol. 30, no. 1, pp. 99–117.

Millikan, RG 1999, 'Historical kinds and the "special sciences"', *Philosophical Studies*, vol. 95, no. 1, pp. 45–65.

Moore, JH 2003, 'The ubiquitous nature of epistasis in determining susceptibility to common human diseases', *Human heredity*, vol. 56, no. 1-3, pp. 73–82.

Okasha, S 2002, 'Darwinian Metaphysics: Species and the Question of Essentialism', *Synthese*, vol. 131, no. 2, pp. 191–213.

Schaffner, KF 2016, *Behaving*, Oxford University Press, New York.

Slater, MH 2015, 'Natural Kindness', *British Journal for the Philosophy of Science*, vol. 66, pp. 374–411.

Sporns, O, Tononi, G & Kötter, R 2005, 'The human connectome: A structural description of the human brain', *PLoS Comput Biol*, vol. 1, no. 4, p. e42.

Tononi, G 2004, 'An information integration theory of consciousness', *BMC neuroscience*, vol. 5, no. 1, p. 42.

Tononi, G & Edelman, GM 1998, 'Consciousness and complexity', *science*, vol. 282,

no. 5395, pp. 1846–1851.

Van Dijk, KR, Hedden, T, Venkataraman, A, Evans, KC, Lazar, SW & Buckner, RL 2010, 'Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization', *Journal of neurophysiology*, vol. 103, no. 1, pp. 297–321.

Watkinson, J, Wang, X, Zheng, T & Anastassiou, D 2008, 'Identification of gene interactions associated with disease from gene expression data using synergy networks', *BMC systems biology*, vol. 2, no. 1, p. 10.

Watts, DJ & Strogatz, SH 1998, 'Collective dynamics of "small-world" networks', *Nature*, vol. 393, no. 6684, pp. 440–442.

Weber, K, Eisman, R, Higgins, S, Morey, L, Patty, A, Tausek, M & Zeng, Z-B 2001, 'An Analysis of Polygenes Affecting Wing Shape on Chromosome 2 in Drosophila melanogaster', *Genetics*, vol. 159, no. 3, pp. 1045–1057.

Weber, K, Eisman, R, Morey, L, Patty, A, Sparks, J, Tausek, M & Zeng, Z-B 1999, 'An Analysis of Polygenes Affecting Wing Shape on Chromosome 3 in Drosophila melanogaster', *Genetics*, vol. 153, no. 2, pp. 773–786.

Williams, PL & Beer, RD 2010, 'Nonnegative decomposition of multivariate information', *arXiv preprint arXiv:1004.2515*.