Kinds, Composition, and the Identification Problem

MSc Thesis *(Afstudeerscriptie)*

written by

Elbert J. Booij
(born January 8th, 1963 in Baarn, Netherlands)

under the supervision of **Dr Ing Robert van Rooij**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
August 28, 2013

Members of the Thesis Committee:
Dr Maria Aloni
Dr Paul Dekker
Stefan Pliquett MA
Dr Ing Robert van Rooij

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Kinds—also known as *natural sets* or *universals*—are a very intuitive assumption about the way the world is put together. As a piece of metaphysical theory, however, they give rise to the *Identification Problem*: which of all sets are the ones that in fact qualify as kinds? In this thesis an answer is given starting out from the assumption that kindhood always coincides with similarity. From this it follows that similarity must be similarity *with respect to*, and properties—kinds—must be arranged in *similarity systems*.

To turn this insight into a credible answer to the Identification Problem, however, a wide variety of (physical) objects must be considered, whose common ground is that they are all *mereologically complex*. Therefore in the second part of the thesis the focus will be on the derivation of *composite kinds*, thus allowing the classification of larger objects in terms of kinds. It will be concluded that classical (Boolean) mereology is sufficient for this purpose. I shall argue that this approach is therefore preferable to that whereby kinds are reified to be (structural) universals.

Contents

Chapter 1

Introduction

Sets play it fair. If we seek to categorize the entities that inhabit our world then, by its own unbiased nature, set theory has it that every collection of entities in full equality is worth to be called a set. This assumption, the result of keeping all options open, leaves a vast space of homogeneity, reminiscent of the description in the famous final scene from Derek Jarman's film *Wittgenstein*:

> *"A world purged of imperfection and indeterminacy. Countless acres of gleaming ice stretching to the horizon."*

This is the default supposition, securing that no unwarranted prior structure is imposed on the domain of interest.

Nature itself, however, is prejudiced: it has strong preferences for this collection of things over that, preferences that make themselves known by the patterns and regularities found in the world, the fact that some conditions entail others, while yet others are excluded; and by our human tendency to classify along certain lines, to perceive similarity here and contrast there, to make this distinction, but not that. *Kinds*, or *universals*, are metaphysical terms for such 'natural' sets. David Lewis (e.g. 1986b) spoke of *properties*, which could be *sparse*, as opposed *abundant*. Sparse properties, universals, and kinds, are all terms meant to capture the same notion, albeit not necessarily by the same theoretical framework.

Among the innumerable possible collections of entities only a very small minority belongs to these natural sets. Whoever takes kinds seriously therefore has to face up to the question of how to identify them. How are we to know if a certain well-defined collection of items is a kind or not? It is by no means generally agreed upon that logic—or even science—has any business here and in the analytic tradition there is even a certain hostility towards the subject, as if interest in it betrays a mildly obscurantist attitude. Quine, in one of his famous essays (1969) went as far as to state that:

> In general we can take it as a very special mark of the maturity of a branch of science that it no longer needs an irreducible notion of similarity and kind. (p 137, 138)

The latter concepts are, in Quine's words, *alien* to logic and set theory. And there is no denying that in first order logic, that great cathedral of formal thought, there is nothing to account for the standing out of the putative kinds.

All this is about the relation between logic and metaphysics, which is a delicate one. As Bar-Am (2008) convincingly argues it was *extensionalism*, the treatment of logical terms as being individuated solely by their extension, that freed logic from the metaphysical commitments that had kept it bogged down ever since Aristotle, and made possible the rapid development that followed. According to Bar-Am it was George Boole who was, more than anyone else, responsible for setting this process in motion. Extensionalism replaced Aristotelian essentialism:

> It was by reducing logic to the study of extensions that Boole was able to transcend, by default almost, some of Aristotle's most stubborn essentialist presuppositions, thereby inaugurating a new era in the study of formal logic. More specifically, Boole succeeded in separating the study of valid inferences from the Aristotelian endeavor to provide the complete taxonomy of all things (and of all things known) that he conflated as both logic and science. (p xii)

The result of this was a revolution in both logic and mathematics of unprecedented magnitude. But there was a second result. In due time a deep suspicion got foothold against metaphysical ideas that had the ambition to 'know better' than mere logic. There were external factors pushing in the same direction, especially the revolutions in empirical science around the turn of the century. Extensionalism, the doctrine that had proved so fruitful to logic, also became the only game in town with respect to metaphysics—to the extent that anyone was still interested in metaphysics in the first place!

As with most positions seeking the extreme of a scale, there is room for qualification. There are good reasons to view the attitude as exemplified by the above quotation from Quine as the manifestation of a misconception which is the mirror-image of that observed by Bar-Am. The emancipation of logic from metaphysics, which made logic thrive, has not been accompanied by a parallel emancipation of metaphysics from logic. There was, so to speak, a changing of roles between master and slave, but otherwise the same entanglement as before. For although it may be true that kinds and similarity are alien to first order logic, they are certainly not alien to science—which is a mild way to put it. If a mature science, as Quine claims, no longer needs these basic concepts, then this maturation is quite a dramatic process, making the thing undergoing it totally unrecognisable from what it was before. The branch of science having reached this final stage would no longer consider *this* way of categorizing objects more salient than *that* (e.g. the class of all electrons v. the class of all electrons plus the Queen's favourite horse). It would effectively mean that those working in the field would *stop classifying*! But classifying is not only the very first thing that every branch of science starts with, but also its ultimate goal. There is a clear sense in which all the activity of science reduces to replacing existing classifications with better ones, where 'better' means: more basic, more all-encompassing, more projectible, with sharper boundaries, and fewer of them.

There can be no doubt that Quine was well aware of all this. He generously and rather exhaustively discusses what many reasons there are to accept kinds, convincingly showing that science is all about refining the concepts of kind and similarity, only to reach, in the very last paragraph, the above-cited conclusion that they will eventually go away—a *non sequitur* so spectacular that it is almost heroic.

This should not, however, blind us to the real philosophical difficulties bound up with the idea that some sets of entities have metaphysical priority over others. They are exemplified by cases like Goodman's *grue*, Quine's *ontological relativity*, and *Putnam's Paradox*. All narratives of this sort have as their common core the claim that such priority has no ground, and can have no ground. This is a much debated issue and I will focus on some aspects only. As I see it, there are two main issues involved. The first is what Lewis calls *abundance* of properties: the idea that every set of entities in the domain of interest determines a property. The second is what Putnam (1980) ironically dubs *metaphysical glue*: the (obscure) principle whereby words or other symbols are linked to their extensions.

Quine's (1969) ontological relativity is mainly about communication, about *agreement* as regards meaning. This does not necessarily involve abundance of properties. One could admit kinds, and still claim that every matching of vocabulary to kinds has several possibilities of recombinant matching that work just as well and are empirically indistinguishable. Putnam's paradox is a far more radical way to put the problem. Via a detour along the Zermelo-Fränkel axioms and Löwenheim-Skolem's Theorem Putnam (*ibid.*) notices that the fact that axiomatic systems in logic may have *unintended interpretations* brings trouble for the Philosophy of Language as well. The ZF-axioms have models that, *contra* common wisdom and intuition, are denumerable. This is certainly not what we *mean* by those axioms: the set of real numbers is famously of higher cardinality. But these countable models have a slightly perverse flavour. The 1-1 correspondence figuring in them is in fact a stronger condition and therefore countability is not 'real' countability, but something formally analogous; hence the 'unintended' status of such models. But then, Putnam asks, if axioms cannot fix a reference, what else could? This problem easily translates to the relation between natural languages and the physical world. Language cannot fix reference by itself. No metaphysical glue! What is needed is something extra: an *interpretation*. But what mysterious sort of object could that be? Is it possible to provide such an interpretation in naturalistic terms, without embarking on radical verificationism ("as long as the truth of sentences can be tested, no one cares about their 'true' reference!")?

Putnam's own reply is in the deeply Wittgensteinian mood that has haunted much of Twentieth Century analytic philosophy:

> To speak as if *this* were my problem, "I know how to use my language, but, now, how shall I single out an interpretation?" is to speak nonsense. Either the use *already* fixes the "interpretation" or *nothing* can.
>
> Nor do "causal theories of reference", etc., help. Basically, trying to get out of this predicament by *these* means is hoping that the *world* will pick one definite extension for each of our terms even if we cannot. But the world does not pick models or interpret languages. *We* interpret our languages or *nothing* does. (p 482)

The problem is real enough, but this solution will only please the desperate. For what could be so special about '*us*' that *we* are able to do things that 'the world' makes no room for—as if we hover above the world, as outsiders? Really naturalistically-minded philosophers should eschew such recourse to exceptionalism. If *we* interpret our languages then it must first be among the features of the world that our languages *can* be interpreted.

The existence of kinds seems to be decisive in providing this condition. Sparseness of properties does not *guarantee* the working of metaphysical glue, even though abundance plainly guarantees its failure. But it does make a difference. If the number of extensions eligible for being named is limited, then the opportunities for alternative word-kind matchings that remain hidden shrink significantly. An even more important additional circumstance is the *correlation* of kinds. Their extensions tend to be related to one another in 'nice' ways, like inclusion (every tree is a plant) and complementary union (every human being is either male or female). This, as I see it, is the background of the often-quoted significance of kinds for the understanding of *causal* relationships. To word such relations we need kinds.

But our words themselves are subject to them as well! Our use of language is behaviour and, as such, embedded in the causal scheme of things. Our ability to categorize, to distinguish between one pair of objects, but not between another; thus using different words in the first, but the same, in the second instance, it part of this scheme. The use of language, the very intelligibility of words and signals, is there only thanks to the fact that the world is no 'gray goo', but has *structure*. This structure was not made by language; it is what makes language possible. To explain *why* it is there, is a challenge of the same order of magnitude as that of explaining why there is something rather than nothing. But to describe it is a task that is not necessarily beyond us.

The most pressing problem with respect to kinds, to my conviction, is what I shall call the *Identification Problem*. Is there a *systematic*, rather than impressionistic, way to identify extensions of kinds? Specifically, given a suggested kind, is there a procedure, at least *in principle*:

(1) to tell whether it really *is* a kind, and

(2) to find out, of a given object, if it *belongs* to it?

Notice that for abundant properties no identification problem arises, since every extension is as good as any other. Anything goes! Not so with kinds. If some sets are 'more equal' than others, then there is something to learn: Which are the ones?

It is easy to underestimate the full scale of the query. Electrons constitute a kind, and so do lumps of gold—fair enough. So long as we stick to small particles (or amounts of them) the choices are not hard to make. But do not pebbles constitute a kind as well? Do octagons, tigers, neutron stars? Consider the cup and saucer before me: are they two of a kind? And what about Ferrari's, or Ecuadorians? It can be legitimately asked if we need kinds for all of these cases, but scaling up from quarks to atoms, molecules, fibres, tissues, etc., there does not seem to be an *obvious* way to draw the line. Almost everything we would like to classify is made of smaller parts that determine its nature. If any meaningful answer to the Identification Problem is to get off the ground, it must include an account of how parts of some kind together make a whole of some kind, by some kind of coalescence. Therefore my theorizing will culminate in some notions about *composite* kinds. To clear the ground for this undertaking, however, we must first face the problem on a more basic level. What is it for a set of objects to form a kind, composite or otherwise?

Some authors, notably Armstrong (1978, 1997), turn to *universals* for an answer. Clearly my use of the term 'kind' is directed towards the same phenomenon, yet my strategy will not be to view, as Armstrong does, universals as real entities, doing the identification by being somehow, self-identically, 'in' the different objects they instantiate. Even though I prefer my ontology to abstain from stipulations explicitly outlawing such constructions, I shall not use them for explanatory purposes. On the contrary: I will show that, especially in the case of complex entities, the theoretical problems that arise with their—structural—universals should be blamed on taking the latter to be real objects, rather than a *façon de parler*. Instead of inventing new entities to order existing ones it is better, to my mind, to focus on what we know about the the latter themselves. My kitchen knife is *sharp* and therefore it can slice a cucumber easily. Although sharpness is a perfect universal, this fact does nothing to make the knife's being sharp anything separable from the knife as it is all by itself. Being sharp is what *the knife* is, not what some other entity is, and it is independent of other things's being sharp as well.

Yet 'sharp' is also a *word*; one which applies to all those sharp entities. It is our human way to express a piece of information about the knife. We have many ways to express such knowledge. Apart from words we have gestures, signals, symbols. We have natural, but also formal, language. All such tools I shall refer to as *concepts*: things, tokens, objects, states of affairs, *in their role of* representing a class of entities. I shall not here make any attempt to account for this phenomenon metaphysically (much of the Philosophy of Mind is dedicated to bring some light in these areas, so I gladly pass the buck). Suffice it to say that it is not only, not even primarily, mental concepts that I want to consider here. I take it as a given that symbols of various nature, for us humans, have the power of representing.

They often do so in a strongly *systematic* way. Pictures, scale models of buildings, Arabic numbers, colour samples, are all vehicles of representation that need only a small amount of precognition to exert unlimited expressive power. In this they differ from e.g. the system of personal identification numbers, that needs a huge database to function. Concepts have an easy-going way of matching their denotations; they can be reused indefinitely, without new information. And concepts are *one-over-many*! In fact I know of no other, non-Platonic, ontologically irreproachable, category of things that have this feature. Concepts are by far the most easy-to-handle access to the phenomenon of universality that we have. Even though we do not quite know *how* they work, we use them on a daily basis.

Concepts do not always express kinds, but for the kinds that we humans can handle we typically have concepts, so it is possible to envisage a language that works with kind-concepts only. In fact this is how I will proceed: I shall develop a formalism, to be called Φ, designed to do just that. Despite the misgivings I expressed about metaphysics being equated to logic, I am strongly committed to logic as a tool for doing metaphysics—and maybe the word 'tool' underdescribes the full significance of this part of the endeavour. Metaphysicians, by the very nature of their occupation, must make do without the stern tutor of empirical observation. Besides common sense logic is all they have by way of an objective guide. So Φ will be a full-blown formalism, providing a certain standardisation of the conceptual apparatus, thus making it easier to sharpen our intuitions about kinds, and to reflect on them. This will help us to display what I take to be the heart of the matter, viz. the relations and correlations between kinds, and subsequently the way complex kinds emerge out of simpler kinds. We shall need this if we want to solve the Identification Problem not only for simple entities, but also—in principle at least—for cats, houses, and lumps of gold.

This notion of concept, therefore, will be my first gateway to understanding kinds. The second is *similarity*. Lewis (1986b) uses this term when he explains what it is for a property to be (abundant or) sparse:

> The abundant properties may be as extrinsic, as gruesomely gerrymandered, as miscellaneously disjunctive, as you please. (...) Sharing of them has nothing to do with similarity. (p 59)
>
> The sparse properties are another story. Sharing of them makes for qualitative similarity (...). (p 60)

Even for hard-boiled extensionalists similarity is a strong intuition, stronger, it seems, than that of kind *per se* (this is remarkable: if it is really true that the extensions of sparse properties are kept together by similarity, then the case for sparse properties is as hard as that of similarity). Carnap's *resemblance nominalism* was an attempt to *reduce* kindhood to similarity relations, later shown, by Nelson Goodman (1951), to be logically inadequate. Rodriguez-Pereyra (2002) has more recently presented a new version of resemblance nominalism, where similarity does not only hold between individuals, but also between pairs of them, and pairs of pairs, etc.

My strong reservations against pictures as advocated by Carnap and Rodriguez-Pereyra do not concern their logical merits; nor am I in any way unsympathetic towards the nominalistic orientation motivating them. My problem is with the utter implausibility of 'reducing' obviously local properties to similarity-relations. That my kitchen knife can slice the afore-mentioned cucumber it does *not* owe to its being similar to other things, far less to its being one of a pair that is similar to other such pairs; but only to its shape and material. Even if resemblance nominalism should come out right logically, metaphysically it turns the world upside down. Two sharp knives are not sharp because they are similar, but similar because they are sharp!

This is not to say that similarity is instead reducible to kindhood. Superficially it seems fine to hold that x is similar to y iff both have *a property in common*. Yet, on closer inspection, this is unlikely to work, the briefest objection being that it is hard to find two objects that are *not* similar under this construal. Most pairs of objects, as wildly unalike as you please, share at least one property, even if trivialities like being self-identical are left out. The discriminatory power of such a criterion falls far short of what we are looking for. A more subtle objection is that similarity, thus understood, is a relation that is reflexive and symmetric, but not transitive. If x is similar to y (because both are blue), and y, to z (because both are a fish), then it is still possible that x and z are not similar (because x is the sky and z is a goldfish). This, I think, squares badly with our intuitive understanding of what similarity is.

So what do kinds and similarity have to do with one another? Al least this: for both we have concepts. 'Similarity' is a concept just like 'red' and 'goldfish', only a relational one. Just like single entities may instantiate redness and goldfishhood, pairs of them may instantiate similarity. As mentioned before, what is characteristic of those special sets that we call kinds is that there are *correlations* between them. In this case the correlation is particularly straightforward: instantiation of the same kind by a pair of entities will correlate with instantiation of similarity.

This, then, will be the central assumption underpinning my theory about identification: kindhood and similarity invariably go together. *No kind without similarity, no similarity without kind!* This may sound like a truism: similarity and kindhood are often viewed as trivially interchangeable. Yet I shall argue that the relation is really an interesting one, and that appreciating it correctly leads much of the way to solving the Identification Problem. I will also show that, given the said assumption, similarity must be *similarity with respect to*. There cannot be one sort of similarity doing all the work that has to be done.

A sound understanding of kindhood and similarity alone, however, will not help us to explain how these categorization principles apply to everyday objects, and why. Most of the objects we are acquainted with are *complex*, they are composed of parts. Now if we know of such a composite entity what kind of entities its parts are, *and* in what kind of way these parts are put together, is there much more that we could wish to know about it—at least in those terms, at least locally?

The question is more than rhetoric. Answering it wholeheartedly in the positive implies a choice for *reductionism*: gaining insight into what kind the whole belongs to proceeds by gaining insight into its parts, preferably down to the very atoms. Denying that this is so implies keeping open the possibility of *emergent* kinds appearing on (ever) higher levels. Here I shall investigate the reductionist option. There are at least two reasons to do so. First, the best ground in general—also for the sciences—to prefer reductionism to any of the alternatives is not that we know it to be true (which we do not), but that reductionism allows the construction of theories with at least some deductive force; theories that are more than assemblies of well-worded intuitions, and whose unforeseen consequences might as well throw our intuitions over. And even though we cannot say that reductionism is true, we know by experience that it is extremely useful, since it is often true to a most satisfactory approximation. Second, even though there might be *more* to the properties of a given whole than only what it inherits from the parts it is composed of, it is hardly conceivable that there is *less* to it. A whole *might* have emergent properties, but it *must* have composite properties!

If this is so, then the solution of the Identification Problem comes down to analysing concepts up to any point where the kindhood of the constituent parts is of a *simple* nature. What I shall demonstrate is that the existence of composite kinds can be indeed be deduced from that of simple kinds, and that this deduction is not trivial. But it can be done making use of *classical* mereology only, and according to this line of reasoning composite kinds are in no crucial way different from other kinds. They behave in the same way, and have the same sort of similarity-relations keeping the things falling under them together. Thus I will argue that *structural universals*, as understood by realists like Armstrong, are of little additional value. We can solve the Identification Problem without them, and they bring problems of their own, as the discussion about methane as a structural universal bears witness to (Lewis, 1986a; Armstrong, 1986; Pagès, 2002; Kalhat, 2008; Mormann, 2010, 2012; Hawley, 2010). These are problems that should not be there; that I take to be an artefact of the theory. There is no straightforward way even to reduce structural universals neatly to their constituting universals (e.g. methane, to carbon and hydrogen), which does not involve what Lewis calls 'magic', or at least non-classical—and, to my mind, quite unnatural—versions of mereology.

This thesis is built up in the following way. In Chapter 2 I shall explain what I take to be the relation between kinds an similarity. I shall make use of the formalism Φ, a modified first order logic, to be presented in Chapter 3. In Chapter 4 a system of (classical) mereology in Φ will be given and in Chapter **??** the all these building-blocks will be put together to arrive at an understanding of composite kinds: kinds instantiated by mereologically complex entities.

Chapter 2

Similarity

2.1 Introduction

My claim has been that the structure of the world is such that kinds and similarity are companions. For every kind there must be some sort of similarity, and where similarity manifests itself, there must be a kind. I have said some words about the relation between the two and here I want to summarize them by the following conditions for a credible concept of similarity:

(1) Similarity is a relation between two entities, coinciding with their being of the *same kind*.

(2) Similarity is an *equivalence* relation.

(3) Similarity is an *abstraction*.

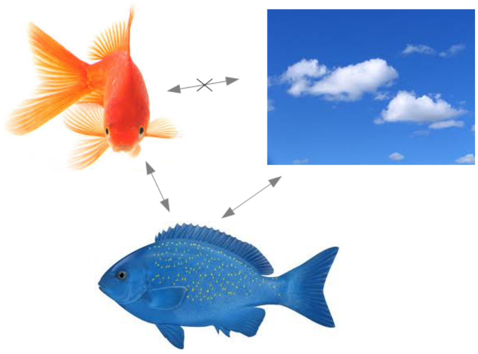
Condition (1) is hardly open to dispute: the being similar of two entities has to do with their sharing some natural property. So, it could be suggested, why not keep things as simple as that: two entities are similar *iff* they share some natural property? Let us call this idea of similarity *E-similarity*.

In the next chapter I shall develop a formal notation especially suited for working with kinds and as a matter of convenience I already want to use some of it here. The first convention with respect to this notation is this: predications like $P(x)$ will be used *only* if the predicate symbol P represents a kind, a natural property. So my use of $Sim(x, y)$ for similarity between entities x and y already betrays my taking this relation to be natural as well—a position I do not expect many complaints about. I will furthermore use the symbol $\Rightarrow$ for inferential relations. In this symbolism, therefore, E-similarity is similarity such that, for any P, the following inference holds:

$$P(x) \cdot P(y) \Rightarrow Sim(x, y)$$

Condition (2) is at the heart of the claim we are about to work out. If similarity is tied to kindhood, then it had better be an equivalence relation; shared membership of a given kind evidently is. Condition (3) is less straightforward, but equally important. Abstractions, as I want to understand the term, are kinds that unify a set of 'underlying' kinds, without being reducible to one of them. Abstraction is never *trivial*: there can be no 'similarity in respect of being red' that all and only means that the two relata are both red. If there is a similarity that two entities share thanks to both of them being red, it should be possible,

Figure 2.1: E-similarity



at least in principle, that two entities be similar *in the same sense*, but thanks to *other* properties. (Their being both blue is an option that comes to mind.)

So can we arrive at an understanding of similarity for which these three conditions hold simultaneously? Applying (1) in the most simple-minded way yields E-similarity. Unfortunately this construal will, in most models, violate (2): it gives us a relation that is reflexive and symmetric, but not transitive. If also $Sim(y, z)$, it might be due to a *different* property that y and z share, e.g. $Q(y) \cdot Q(z)$. But then why should x and z have to share a property?

This might lead one to cast doubt on (2), but I think this doubt is unwarranted. Although, as the argument just given shows, (2) is a substantial claim, it is not a particularly venturous one. Similarity and equivalence are very closely related concepts—without much exaggeration one could say that the formal notion of equivalence has been invented to cover the intuitions behind similarity. Transitivity is an essential part of them. I feel honoured to have someone of the stature of Euclid to my witness, whose Book of Elements, Common Notions, opens with exactly this maxim:

> Things which equal the same thing also equal one another.

Surely there are similarity-like concepts that are not transitive: *near*-similarity for instance. Yet if we take a closer look at the various versions of near-similarity we find no reason to relinquish the demand for transitivity. First, near-similarity is 'near-transitive': If x looks a lot like y, and y, a lot like z, then x cannot look *too* different from z (pulling x towards y, and y towards z, is somehow also pulling x towards z). Second, it is slightly odd to have a concept of near-similarity without the (in all likelihood *simpler*) one of exact similarity. At the very least it should be there as a limiting case.

We could of course tinker with our models a bit to enforce transitivity. Given $P(x) \cdot P(y)$, the way to prevent $Q(y) \cdot Q(z)$ $(Q \neq P)$ is to prevent $Q(y)$, and a slightly odd but instructive way to do this would be to demand that every individual necessarily have *one* property at most. A more convincing variation on this design is obtained by switching to a new set of properties, each of whom is really the *conjunction* of every property (in the old sense) an item has. Clearly of this new set everything has only one property. Call this *A-similarity*; it produces a beautiful equivalence-relation: a partitioning of everything into classes of indistinguishable entities. Of course A-similarity is too limited; yet every credible set-up of similarity will *also* include the borders that A-similarity draws.

The alternative is to tinker with the concept of similarity itself. If we do not want to let $Q(y) \cdot Q(z)$ spoil the similarity thanks to $P(x) \cdot P(y)$, we could say that there is *more than one type* of similarity. A very natural conception giving us just that is similarity *with* *respect to* this or that. The obvious first attempt should be to distinguish similarity with respect to *any* of the original predicates. So $Sim_P(x, y)$ will mean that either both x and y are P, or none of them are.

This proposal, however, still has a fatal flaw. In most models it violates condition (1) since it makes things similar that have *nothing* in common. If x is not P and neither is y (and both have nothing in common otherwise), then they are nevertheless similar in respect of P! Step two, therefore, could be that, for every original predicate P, its *complement*, $\overline{P}$, is introduced as a predicate as well. Then $Sim_P(x, y)$ entails that either $P(x)$ and $P(y)$, or $\overline{P}(x)$ and $\overline{P}(y)$. Notice that, instead of Sim_P, we could as well write $Sim_{\overline{P}}$ now; they are the same. Better still, to avoid biases, we could write $Sim_{\mathbf{P}}$, with $\mathbf{P} = \{P, \overline{P}\}$.

This, I think, begins to look like an acceptable notion of similarity. It produces equivalence relations and, by conjunction, also a host of new similarity relations, each as good as the original ones, including (by conjunction of all of them) A-similarity.

2.2 Saturation

This has been achieved thanks to the fact that we have adjusted the models so as to make them *saturated* with respect to their original predicates. I shall use this term to indicate the condition that for every predicate there is an *alternative*. In this case, whatever is not P is $\overline{P}$ and nothing is both P and $\overline{P}$; and this goes for every predicate.

An element, however, has been slipping into the argument the impact of which should not be taken lightly: the assumption of *negative* predicates. This decision has two repercussions. On the metaphysical level it means that, since predicates stand for kinds, we have committed ourselves to the existence of negative *kinds* as well, a famous non-starter according to most of those studying the matter; I shall return to this issue shortly. On the logical level these predicates, at first glance, look perfectly well-behaved. They are not. The troubles surface immediately once we consider *combined* similarities, like similarity with respect to P and Q. If we accept conjunctive kinds—as most authors happily do—this should be unproblematic: if P and Q are original predicates, then the predicate $P \cdot Q$ must be admitted also. Following the recipe described above, similarity in respect of $P \cdot Q$ becomes $Sim_{\mathbf{R}}$, with $\mathbf{R} = \{P \cdot Q, \overline{P \cdot Q}\}$. On the other hand, taking into account that similarity with respect to P is the same thing as similarity with respect to $\overline{P}$, analogous for Q, it would be strange if similarity with respect to $P \cdot Q$ would differ from similarity with respect to $\overline{P} \cdot \overline{Q}$. But $Sim_{\mathbf{R}'}$ ($\mathbf{R}' = \{\overline{P} \cdot \overline{Q}, \overline{\overline{P} \cdot \overline{Q}}\}$) is certainly *not* the same as $Sim_{\mathbf{R}}$.

Surely we *could* opt for multiple similarities in case of combinations of predicates (we would have four of them, in this example) but the intuitive credibility of this arrangement is low. There is little doubt that the similarity we are looking for is a simple conjunction of $Sim_{\mathbf{P}}$ and $Sim_{\mathbf{Q}}$. Now this conjunction is $Sim_{\mathbf{R}''}$, with $\mathbf{R}'' = \{P \cdot Q, P \cdot \overline{Q}, \overline{P} \cdot Q, \overline{P} \cdot \overline{Q}\}$. Instead of one complement $\overline{P \cdot Q}$ we have *three* distinct alternatives filling up the space around $P \cdot Q$.

This, as I see it, brings the following new insight. Similarities like $Sim_{\mathbf{P}}$, as construed in the last chapter, are similarities with respect to *binary* conditions, but this is by no means essential. Instead of stipulating that every P be complemented with its own $\overline{P}$, we could make the more general demand that the original predicates come in *sets of alternatives* such that every individual must instantiate *exactly one* predicate of the set. The cardinality of this set can be two, but equally well any bigger finite or infinite number. A set of this sort plus its similarity relation will be called a *similarity system*: $\mathfrak{S} = \langle \mathbf{P}, Sim_{\mathbf{P}} \rangle$.

Similarity systems are surprisingly natural constructions; once you pay attention, you see them everywhere. With respect to material objects *colour* is a good example: it seems that every (part of) every material object has to be coloured one way or another. Here the number of elements in the set of alternatives is rather large: the human eye can distinguish several millions of different shades of colour. Now colours are very much conditions belonging to the 'manifest image', but in theoretical physics similarity systems are abundant: charge, mass, spin, spacial location, etc. Similarity systems, especially those with a large or infinite number of alternatives, are reminiscent of *dimensions*[1], but maybe a better identification is with *determinables* and *determinates*, as the terms were introduced by W.E. Johnson (1921). Here the similarity system is the determinable and the elements of the set of alternatives, the determinates[2].

There are, apart from the above exposition about similarity, two independent reasons to take saturation seriously. First, suppose we have two objects: A is white, marble, and cubic, and B is marble and cubic only. Can this really be the case? I would say that, if B is not white, it must have a different colour: black, for instance. It seems to be so that you cannot just take away a property without replacing it with something else. The descriptions of A and B are perfectly acceptable *concepts*, with that of A having a higher degree of specificity than that of B. But two *objects* one of which just has one extra property? Assuming complement properties, as we did with the introduction of the $\overline{P}$'s, is the more intuitive line, but then we conjecture the existence of a new, real property—exactly one, moreover. The mere lacking of a property does not *automatically* create its complement as a replacement. Therefore my hunch is that our willingness to accept $\overline{P}$-like properties is really the assumption of saturation.

The second point is epistemic in nature. It is no small riddle that humans can perceive things to have certain properties, but it is not even more puzzling that we can perceive things to *lack* properties? How can we see absence, given that absence is not something itself? The chair a few meters from me is not green. How can I tell? The answer is simple: because it is red! How can I know that there is no rhinoceros in the room? Because I see that every portion of space big enough to be a rhinoceros is transparent, and the stuff such creatures are made of is not. For the same reason I cannot *perceptually* assure myself that there is no ant in the room, or no ghost.

The second argument has a loose end that I will shelve now, but that will have to be dealt with some day. It is all fine that the redness of an object excludes its being green, but how do I know *that*? For if I do not have this sort of information, the argument gets nowhere. This is an extremely delicate matter: it has to do with necessity or with *a priori* knowledge, or both. The issue is deserving of further investigation, but it will not be done in this text.

The important consequence is this. If saturation is assumed, if predicates line up in sets of alternatives, then they can be categorized in *types*. That is what the predicates in the example just given differed in. Types will be represented by the set of predicates belonging to them in brackets e.g. $(\mathbf{P})$. A *concept* will be declared of type $(\mathbf{P})$ iff every argument in it is predicated in terms of $\mathbf{P}$. Concepts can be specific in terms of more than one type, in which case we simply write, e.g. $(\mathbf{PQ})$.

Not every concept has a type. If $\mathbf{P} \neq \mathbf{Q}$ then:

$$P_5(x) \cdot Q_1(x)$$

[1]The association is natural, but the analogy is convincing only where there is a fixed basis. If dimensions tend to 'merge' (cf. colour) it is better to view them as all belonging to one similarity system.

[2]Sensory qualities like pitch and taste are notoriously difficult to fit into this scheme, as different determinates do not seem to be mutually exclusive. Although this is a serious point, it is not as fatal as it is often taken to be. Maybe hearing or tasting is more like watching a picture, where different colours can co-exist in perfect (dis-)harmony. Mereological concepts may be useful to shed light in matters like these.

($P_5 \in \mathbf{P}$, $Q_1 \in \mathbf{Q}$) is of type ($\mathbf{PQ}$), but:

$$P_5(x) \cdot Q_1(y)$$

has no type, since the argument x is not specified in terms of the set of alternatives $\mathbf{Q}$.

2.3 Disjunction and Negation of Kinds

A small but significant step towards solving the Identification Problem is to reach a verdict on whether the set of kinds is closed under conjunction, disjunction, and negation. This is an old conundrum, not to say a threadbare topic in metaphysics. The intuition is that it is indeed closed under conjunction, but not under the other operations. Notice that *infinite* disjunction and (in combination with conjunction) negation wipes out most of the border between sparse and abundant properties, so let us concentrate on the finite case. Let us take two kinds, P and Q, and have a look at the following claims:

(4) $\overline{P}$ (not being P) is a kind.

(5) $P \cdot Q$ (being P and Q) is a kind.

(6) $P \vee Q$ (being P or Q) is a kind.

Which of them is true? If we think extensionally then all three are, maybe with the restriction that the extension should not be empty. If we let our intuition run freely, however, then (5) is true, maybe with the restriction that P and Q should not be contradictory, and (4) has some credibility. But (6) is certainly false. Disjunction is the least popular of kind synthesizers. In Armstrong's (1991) words:

> Consider an a that has property P but lacks Q, while b has Q but lacks P. Is it not a joke to say that they have ($P \vee Q$) in common? (p 197)

This way of pressing the point it is convincing, but only because it is restating a powerful intuition; *not* because there is any deductive force in it. Most arguments against disjunctive kinds suffer from this subtle circularity. The claim is that $P \vee Q$ cannot be a real kind because something that is P and something that is Q need not have anything in common—but this is because we do not accept that kinds are closed under disjunction! A very similar sleight of hand can be seen in Armstrong's (1978) appeal to the causal powers of disjunctive kinds:

> Suppose, again, that a has P but lacks Q. The predicate 'P∨Q' applies to a. Nevertheless, when a acts, it will surely act only in virtue of its being P. Its *being P or Q* will add no power to its arm. This suggests that *being P or Q* is not a property. (p 20)

But causes—of all things—are notorious for their ability to get intractably entangled. If sixty-two bullets hit a man, which of them can be said to have caused his death? The sixty-second bullet, as it seems 'adds no power' to the other ones in the volley. I share Armstrong's unease about causal powers of disjunctive origin, but I do not know how to communicate it to someone who simply claims that being hit by one of those bullets causes death. Or that being sick or poor causes misery.

With the idea of saturation at our disposal, however, we seem to be in the position to forge a strong *logical* argument for the acceptability of conjunctive, but not disjunctive or negative kinds. This is the rough line of argument: if kinds and similarity coincide, thus similarity is an equivalence-relation, then *every* kind must find its place in some similarity system. If we can show that this condition is unproblematic for conjunctive kinds, but at least unlikely to obtain for disjunctive and negative predicates, we have what we want.

If we have a model that is saturated, then for any concept $P(x)$ $(P \in \mathcal{P})$ there will be a similarity system $\mathfrak{S} = \langle \mathbf{P}, Sim_{\mathbf{P}} \rangle$ $(\#\mathbf{P} > 1)$ such that $P \in \mathbf{P}$. For convenience we shall write $P_1, P_2 ...$ for the elements of $\mathbf{P}$. First we shall show that the most straightforward way to construct a similarity system in which the putative new predicates fit works only for conjunction:

- **Negation**. Suppose that for any $P \in \mathcal{P}$ there is a predicate $\overline{P}$ such that $\overline{P}(x)$ iff not $P(x)$. Assume that there is at least one similarity system, let it be $\mathfrak{S}_{\mathbf{P}}$, with $\#\mathbf{P} > 2$ (if there is more than one at least binary system then this condition is already met by their conjunction). By saturation every entity a must instantiate one of the concepts in $\mathbf{P}$, so let us say $P_1(a)$. But then neither $P_2(a)$ nor $P_3(a)$, hence $\overline{P_2}(a)$ and $\overline{P_3}(a)$. Therefore $\mathbf{R} = \{\overline{P} \,|\, P \in \mathbf{P}\}$, as it members $\overline{P_2}$ and $\overline{P_3}$ are not mutually exclusive, does not constitute a set of alternatives for a similarity system.
- **Conjunction**. Suppose that for any $P, Q \in \mathcal{P}$ there is a predicate $P \cdot Q$ such that $P \cdot Q(x)$ iff $P(x)$ and $Q(x)$. Let $\mathbf{R} = \{P \cdot Q \,|\, P \in \mathbf{P} \text{ and } Q \in \mathbf{Q}\}$. Clearly $\mathbf{R}$ is a set of alternatives: by saturation every entity a must instantiate exactly one element of $\mathbf{P}$ and one, of $\mathbf{Q}$, hence $P_i \cdot Q_j(x)$ for exactly one pair i, j. The corresponding relation $Sim_{\mathbf{R}}(x, y)$ is coextensive with $Sim_{\mathbf{P}}(x, y) \cdot Sim_{\mathbf{Q}}(x, y)$. Thus $\mathfrak{S} = \langle \mathbf{R}, Sim_{\mathbf{R}} \rangle$ is a similarity system for any given $P \cdot Q$. (*Remark*: if $\mathbf{P} = \mathbf{Q}$ then the extension of many of the new predicates will be necessarily empty. This does not pose a problem for their existence.)
- **Disjunction**. Suppose that for any $P, Q \in \mathcal{P}$ there is a predicate $P \vee Q$ such that $P \vee Q(x)$ iff $P(x)$ or $Q(x)$. By saturation every entity a must instantiate one of the elements of $\mathbf{P}$, so let us say $P_1(a)$. But then also $P_1 \vee Q_1(a)$ and $P_1 \vee Q_2(a)$. Therefore $\mathbf{R} = \{X \vee Y \,|\, X \in \mathbf{P} \text{ and } Y \in \mathbf{Q}\}$, as it members $P_1 \vee Q_1$ and $P_1 \vee Q_2$ are not mutually exclusive, does not constitute a set of alternatives for a similarity system.

Conjunction preserves the structure of similarity, whereas negation and disjunction do not—at least, under very moderate presuppositions, not in the same straightforward way. Of course it could be objected that there might be a different way to construct a suitable set of alternatives, but the prospects for achieving this look bleak. At the very least it will not be possible to harbour all the $\overline{P_i}$'s or the $P_i \vee Q_j$'s in one and the same system, as can be done with conjunction, since they will typically overlap. So there would have to be different similarity systems: in the case of negation one for every one of them, the complementing predicates being of a different sort. I, for one, have not the slightest idea what such a construction would look like and I fee confident to contend that there is no convincing candidate to be found.

Let us take stock. If the above is correct, then at least some logical flesh has been laid upon the metaphysical bones of the concept of *kind*. If we consider models that start with only predicates for kinds, a plausible way to introduce new kinds is by conjunction of old ones. But the similarity predicates themselves too seem to have every claim to kindhood—as relational kinds this time. For models this means that their set of predicates must have a certain structure to be 'ontologically adequate'. Finally, and most importantly, adopting kinds in this way implies departing from a purely extensional basis for predicates.

2.4 Comparison

There is a natural extension to this line of thought. If saturation is an all-pervading feature of our models, as appears to be dictated by the intertwinedness of kindhood and similarity, then similarity *itself*, being a kind as well, should also be subject to saturation. Let us see what this may look like. Let the similarity in question be $Sim_{\mathbf{P}}$, occurring in the system $\mathfrak{S} = \langle \mathbf{P}, Sim_{\mathbf{P}} \rangle$.

Reiterating the argument for saturation applied to this case: would it not be strange if, when we take two items x and y, a higher-level supervenient relation like $Sim_{\mathbf{P}}(x, y)$ obtains whenever $P_i(x) \cdot P_i(y)$, but nothing of the kind is true when $P_i(x) \cdot P_j(y)$ with *different* i and j? Consider colour again: if two items are not fully similar in this respect, there is definitely more to tell about the way they are related than only that they are *different*. The relation between purple and crimson is clearly distinguishable from that between purple and lemon: the former is closer to similarity than the latter. Other relations, providing alternatives to similarity, but otherwise equally natural, may enter the stage as well.

There is, I think, every reason to believe that saturation also applies to the level of *supervenient* relations. I shall call these relations, the 'noble gas' among which is similarity itself, *comparative* relations. They constitute a very important class of supervenient relations. To illustrate the idea I will start with the simple yet instructive example of elementary charge in physics. Let the similarity system be $\mathfrak{C} = \langle \mathbf{C}, Sim_{\mathbf{C}} \rangle$, with:

$$\mathbf{C} = \{C_1, C_2, C_3\}$$

also known as:

$$\{Pos, Neutr, Neg\}$$

I shall use the notation $Cmp^{i,j}_{\mathbf{C}}(x, y)$ for the comparative relation supervening on $C_i(x) \cdot C_j(y)$. Furthermore I shall write:

$$Comp(\mathbf{C}) = \{Cmp^{i,j}_{\mathbf{C}} \mid i, j \leq \#\mathbf{C}\}$$

In words: given a set of alternatives $\mathbf{C}$ there is a second set of alternatives $Comp(\mathbf{C})$ of supervenient two-place relations, the most prominent of which is the relevant similarity. This can be expressed by means of the following inferences:

(7) $C_i(x) \cdot C_i(y) \Rightarrow Sim_{\mathbf{C}}(x, y) \qquad (i = 1, 2, 3)$

(8) $Pos(x) \cdot Neutr(y) \Rightarrow Cmp^{1,2}_{\mathbf{C}}(x, y)$

(9) $Neutr(x) \cdot Neg(y) \Rightarrow Cmp^{2,3}_{\mathbf{C}}(x, y)$

(10) $Pos(x) \cdot Neg(y) \Rightarrow Cmp^{1,3}_{\mathbf{C}}(x, y)$

We have already taken for granted that $Cmp^{1,1}_{\mathbf{C}}(x, y) = Cmp^{2,2}_{\mathbf{C}}(x, y) = Cmp^{3,3}_{\mathbf{C}}(x, y) = Sim_{\mathbf{C}}(x, y)$. Now physical wisdom has it that the also the step from *Pos* to *Neutr* is the same as that from *Neutr* to *Neg*, so $Cmp^{1,2}_{\mathbf{C}}(x, y) = Cmp^{2,3}_{\mathbf{C}}(x, y)$. To express this we shall also write $Cmp^{a}_{\mathbf{C}}(x, y)$ for this relation, to arrive at the more perspicuous:

(11) $Pos(x) \cdot Neutr(y) \Rightarrow Cmp^{a}_{\mathbf{C}}(x, y)$

(12) $Neutr(x) \cdot Neg(y) \Rightarrow Cmp^{a}_{\mathbf{C}}(x, y)$

(13) $Pos(x) \cdot Neg(y) \Rightarrow Cmp^{b}_{\mathbf{C}}(x, y)$

(The default convention will be that different letters indicate different relations). It seems that this is how the comparative relations are arranged in this domain. And, of course, the following inference holds:

(14) $Cmp^{a}_{\mathbf{C}}(x,y) \cdot Cmp^{a}_{\mathbf{C}}(y,z) \Rightarrow Cmp^{b}_{\mathbf{C}}(x,z)$

Comparative relations are not very exciting unless if we can make such identifications: if we have reasons to declare certain of these $Cmp^{i,j}_{\mathbf{R}}(x,y)$ relations to be *the same* concept. This is where experience comes in. None of the identifications about charge could have been foreseen by the theory of comparative relations only. These are all *physical* facts. They are facts, however, that are embedded within a logical structure.

Comparative relations are archetypical *supervenient* relations. They are also to be viewed as relational *abstractions*—these terms I take to be synonymous. Supervenient relations are relational kinds that are dependent on other 'underlying' kinds. Even though there are inferences both ways, there are reasons to believe that the relation between them is *asymmetric*. Surely the following inference holds:

$$Red(x) \cdot Sim_{col.}(x,y) \Rightarrow Red(y)$$

but it does not seem right to say that y's being red supervenes on x's being red plus the colour-similarity between both. Therefore my hunch is that there are *levels* of kinds, but I shall not work out this idea here.

Chapter 3

The formalism Φ

3.1 Introduction

No logic is mere calculus. Systems of logic reflect the way we the designers believe thought and reasoning to be, and the world to be. For first order logic this is especially obvious, since it brings its own metaphysics in the guise of its semantics. This semantics reveals what first order logic 'really is about': it is about *individuals* and *classes* they may belong to. In its plain form these classes, labelled by *predicates*, may be any set of individuals. In fact they may be any set of anything: other sets, numbers, relations, expressions, etc.; all so long as they can be conceived of *as* individuals. In this 'minimal metaphysics' there are no kinds. Different models may have different sets of predicates, but first order logic tolerates every non-empty set of them, without preferring one to the other. And there need not be similarity either, since similarity, as construed here, is a supervenient predicate, i.e. a predicate tied to the other predicates. In first order logic such a rule 'producing' new predicates is admissible, but by no means standard.

If we want a logic enabling us to speak about kinds, the obvious first thing to do is stipulate that its predicates *refer to kinds only.* To integrate this feature into the structure of the logic, we must secure saturation by introducing sets of predicates supporting complementary extensions. And we must demand of the set of predicates of the model that it contain similarity-predicates and predicates for comparative relations—and many, many others; one for every new kind that will emerge. With minor modifications all of this can be realized within first order logic; it is essentially strong enough for anything that, in my opinion, is needed to describe a world of kinds.

In fact it is unnecessarily strong. The formalism Φ, as I want to develop it here, can be seen as a dialect of first order logic, trimmed down to be able to deal with kinds in a simpler and more perspicuous way. But the adaptations are also grounded in differences in *interpretation.* These are subtle yet significant. In first order logic the main focus is on *individuals*: it is individuals that are being quantified over and that form classes. They may, for modal purposes, even occur in different worlds, with or without trans-world identity. In Φ the focus is on kinds. This does not imply realism: it is not necessary (nor prohibited, for that matter) to reify kinds as 'universals', with a life of their own. They need not be anything over and above their being the way individuals are. But referring to that—to *the way they are*, i.e. their kind—is the only way to refer to these individuals.

This is illustrated by the most basic piece of symbolism in Φ. Just like in first order logic predication is done by a formula in which the predicate and the entity each have their own symbol, playing very different roles:

$$P(x)$$

In Φ the division of labour between these two sorts of symbols should be thought of as carried to the extreme. The predicate-symbol P takes care of whatever is said of the entity-symbol x, which, in its turn, is nothing but a place-holder. It is, as Armstrong would have it, a *thin particular*—as thin as they come. It indicates that there is an entity and it binds predications together in case of repeated use of the same symbol. *Everything* else is in the predicate-symbol, which, on the other hand, signifies nothing *but* a predication. The predication must reveal (an aspect of) what the entity is like. I shall call this *thin predication.*

An interesting consequence of thin predication is that it cannot ever go beyond the entity x. Hence, the following formula would not be admissible in Φ:

$$*\,Father(x)$$

because $Father$—in its most obvious interpretation—would express a condition stretching beyond the person x who is the father. It contains a 'hidden entity'. Thin predication with respect to this condition is only possible if we use a *relational* predicate:

$$FatherOf(x, y)$$

where y stands for the child of what x stands for. Φ cannot express relational properties other than in the form of the relations they are, because monadic predications can only convey *local* properties.

On the other hand, a consequence of the 'thinness' of x in $P(x)$ is that what *it* stands for can never be more than its bare existence. It is entirely empty of predication. The occurrence of x in the formula indicates that there is mention of an entity, plus that it is the same entity as possible x's elsewhere in the same formula. It is a mere variable, keeping predications together by anaphoric bonds. In Φ we will, as a limiting case, admit *null predication*, yielding a *bare formula*, one of (an) argument(s) only. In many expressions bare formulas are indispensable to convey what we want.

Predication, as exemplified by $P(x)$, is the most fundamental bearer of meaning in Φ. This type of expression, the one that conveys *concepts* in Φ, I take to be so basic that I see no justification for the introduction of additional primitive types of referring symbols, like proper names or individual constants of other sorts. Although these belong to the standard furniture of first order logic, from an ontological point of view they have little to their recommendation. If proper names are an exceptional type of *kind*, whose extension is necessarily a singleton, then we do not need special arrangements for them. But as of today this is a relatively unusual way of thinking about names. They are more commonly associated with a direct, or *indexical* way of relating to what they refer to, as opposed to the *descriptive* mode that uses kinds. Saul Kripke's (1981) very influential thoughts on the matter, that have been developed and modified in various ways, have as their core the idea that the referent of a proper name gets determined by an event of e.g. pointing at it: *this* we call Mount McKinley. Afterwards the name remains stuck to the object by a process of the use being transferred from one user to the other, thus referring to *that very object*, not to some description or other. The primordial act of baptising could even have been done by a description ("the huge peak we saw after a two days walk"), but that description is not the meaning of the name. In fact there is no meaning of the name, but its use for that object.

Convincing though Kripke's reasoning be *contra* proper names as (linguistic) descriptions; whether the use of proper names stands completely on its own feet is quite another matter. If I see a mountain that I suddenly realize is Mount McKinley, or if I recognize it on a picture, all I have to go by is still properties. Even with full-blown indexicality at work, as when someone points and says: "Do you see *that*? That is Mount McKinley!", the 'that' reduces to a certain type of spacial relation between the pointing person and the object. But if I am enough an expert on the matter, and someone gives me a detailed description, at some point I may feel confident to decide that what she describes *must* be what I have come to know under the name 'Mount McKinley'. It is not that in any of those examples the meaning of 'Mount McKinley' coincides with one of these modes of presentation; all they do is *convince* us that the object in the distance is *the same as the one that once has been named so*, and that is enough reason for us to use the name.

First order logic commonly mimics this human behaviour by the formal device of individual constants, items that are purportedly 'given' by some epistemic shortcut, bypassing their properties. The practical use of this device for modelling linguistic meaning is beyond doubt, but it is highly questionable, or so I contend, that anything corresponds with it metaphysically. Jubien (1993), in this respect, speaks of the *Fallacy of Reference*:

> This tendency [= to confuse 'is' of predication with 'is' of identity] is really just a symptom of a deeper disorder in our normal philosophical thinking. That disorder is the belief that ordinary proper names and at least some definite descriptions actually *refer to* (of *denote*, or *designate*) specific entities. (p. 22)

Kripke is right insofar that a name is intended to stand for some object, not for any specific description of it. But neither at the act of baptising, nor anywhere along the causal chain of uses of the name, was there a magical 'direct' semantic access to the object so named. It was given to us by our senses, by descriptions, or by our inferences and conceptualizations. *Haeceities*, as they are often called, plus all that comes with them, including names, are among our most persistent intuitions, but the least that can be said is that, for our present purposes, it is more parsimonious to leave them out.

3.2 Inferences

The most important type of formula in Φ will have the following general meaning:

> *Given* a state of affairs such and such, *it can be concluded* that this and that is the case with it.

Here the antecedent is the assumption of the existence of an entity—or a group of them—under a certain *positive* description. It is a predication, or a conjunction of them. The consequent also consists of (a) predication(s), but they can be in a negative or a disjunctive position. A formula of this sort will be called an *inferential sentence*. Inferential sentences serve to express the basic relations between kinds. Rough-and-ready examples are:

(15) $Black(x) \Rightarrow \neg White(x)$

(16) $Square(x) \Rightarrow Rectangle(x)$

(17) $Human(x) \Rightarrow Man(x) + Woman(x)$

(where + stands for exclusive disjunction; below I shall spend some extra words on the peculiarities of this connective).

The inferential arrow (⇒) is the pivot of the sentence. The subformula preceding it, the antecedent, is what we shall call a *conceptual formula*: an atomic formula or a conjunction of atomic formulas. Conceptual formulas are so called because they are the expressions in Φ that stand for concepts. They can be simple predications, but also highly complex conjunctions of relational and monadic predications. But what they express is always a positive type state of affairs, free of negation or disjunction.

The subformula behind the inferential arrow, the *consequent*, is a *logical formula*: a propositional function of conceptual formulas. This will allow the inferential sentence to express the basic relations between kinds:

(18) Whatever is black, cannot be white.

(19) Whatever is a square, must be a rectangle.

(20) Whatever is a human, must be either a man, or a woman.

What is the nature of the inference? As the examples suggest, the relation should be interpreted as being *fully general*: its validity means that it holds always, where 'always' includes every situation that *could* have obtained, even if it does not actually. The customary way of handling this logically is to say that the consequent holds in every possible world where the antecedent does. This is surely correct; observe, however, that, since the antecedent is a *type state of affairs* rather than, as is more commonly the case in logical systems, a fact, it is repeatable also in the actual world. In case anyone would not be prepared to suffer the ontological hardships of possible worlds, and therefore try to base her notion of validity on the actual world only, she would already get much of what she needed. All actual black things *are* not white, after all! A setback for this plan would of course be the contingency of there not being an instantiation of the antecedent, or of the consequent holding generally by mere coincidence.

This would make such an actualist modality unsatisfactory. However, or so I claim, it is unsatisfactory because it is fallible, sloppy—*not* because it is wrong-headed in principle. For what do we in fact know about non-actual worlds, other than what can be concluded from extrapolation from what is the case in the actual world? And the scheme can easily be improved on. Take this example:

(21) $Black(x) \cdot Unicorn(x) \Rightarrow White(x)$

Even if there happens to be no black unicorn in the actual world, we still feel that the inference should be called invalid. Why? Because we do not want things that just 'happen' to be so, to interfere with the validity of inferences.

There is a way around this problem. We could base our notion of validity on *ignorance* about contingencies. This means that we embark on an *epistemic* interpretation of modality. Suppose one could be sure about the correctness of certain *basic inferences* in the actual world, but were unknowing about everything else that is the case there. Then, for any inference, validity would mean that an intelligent thinker can be sure of its (actual) correctness. If so, since this thinker would not be able to exclude the existence of unicorns, black or otherwise, the vacuous fulfilment of an inference like (21) would be blocked. Clearly this recipe also works against the fulfilment of inferences by coincidental truth of the consequent. As we will indeed arrange things so that, in Φ, all valid inferences can be traced to a limited number of basic ones, this understanding of modality is tailor-made for our purposes.

This epistemic interpretation of modality will not lead to different results, but it sheds a slightly different light on what possible worlds really are. Construed in the way described they are not so much different *worlds*, but different *ways* the actual world could be for someone with incomplete knowledge. But it is still this world—which in fact *is* unknown to us in many respects. Possibilities are, *contra* Lewis, just that.

Consequently in Φ there is no 'brute' generality. Those generalities that do obtain can only follow from the instantiation of kinds. If ravens are always black there must be something about ravenhood that necessitates blackness, otherwise the generality is mere coincidence and, from the perspective of Φ, not interesting enough to express. A separate way, therefore, to express generality that does not arise from relations between kinds is omitted in Φ: there is no explicit general quantification (∀). The inferential arrow, however, expresses an implicit type of general quantification. (There is no explicit existentional quantification either, but see below.)

What these considerations about generality show is that kinds and modality are two sides of the same coin. Modality is not an extension of the ideas about kinds; kinds *precede* modality. Take the cup before me. It is blue, but it could have been yellow. Had it not been blue, the set of blue things would not have been the set that it actually is, but the *kind* blue would have been exactly the same kind. It *must* be, for how otherwise could we make sense of the facts in our alternative world, viz. that the cup is yellow, and not blue?

3.3 Syntax and Semantics

After many preliminary remarks, this section will be devoted to a precise description of Φ. A language $\mathcal{L}$ of Φ will consist of the following symbols:

- a set $\mathcal{P}$ of *predicate symbols*
- a set $\mathcal{X}$ of *entity symbols*
- the set $\{\neg, \cdot, \vee, +, \bot\}$ of *logical constants*
- the *inferential arrow*: $\Rightarrow$
- the *realizability symbol*: $\Diamond$

Additionally we shall make use of the following notations:

- $\bigodot_{i \leq n} \phi_i$ is the generalized conjunction $\phi_1 \cdot \phi_2 \cdot ...\phi_n$.
- $\bigoplus_{i \leq n} \phi_i$ is the generalized exclusive disjunction $\phi_1 + \phi_2 + ...\phi_n$.
- $arg(\phi)$ is the set of arguments occurring in the formula ϕ.
- $\mathbf{x}$ is the array $\langle x_1, x_2, ..., x_n \rangle$.
- $R(\mathbf{x})$ is the formula $R(x_1, x_2, ..., x_n)$.

The option that $n = 0$, thus $\mathbf{x}$ is a null-array, is included (not, however, that R has zero arguments!). We shall often treat $\mathbf{x}$ as a set, yet $\#\mathbf{x}$ indicates the length of the array (n), not the number of different elements.

In Φ we shall take the following formulas to be well-formed:

Definition 3.1. *conceptual formula*

- If $\phi = x$, with $x \in \mathcal{X}$, then ϕ is a conceptual formula.
- If $\phi = R(\mathbf{x})$, with $R \in \mathcal{P}$ and $\mathbf{x} \in \mathcal{X}^n$, then ϕ is a conceptual formula.
- If $\phi = \psi \cdot \chi$, with ψ and χ conceptual formulas, then ϕ is a conceptual formula.

formulas of the form $\bigodot_{i\leq n} x_i$, with $x_i \in \mathcal{X}$, will be called *bare formulas*. We shall take the liberty to abbreviate them also by $\mathbf{x}$, as the context will disambiguate.

Definition 3.2. *logical formula*

- If ϕ is a conceptual formula, then ϕ is a logical formula.
- If $\phi = \neg\psi$, where ψ is a logical formula, then ϕ is a logical formula.
- If $\phi = (\psi \cdot \chi)$, where ψ and χ are logical formulas, then ϕ is a logical formula.
- If $\phi = (\psi \vee \chi)$, where ψ and χ are logical formulas, then ϕ is a logical formula.
- If $\phi = (\bigoplus_{i\leq n} \psi_i)$, where, for each $i < n$, ψ_i is a logical formula, then ϕ is a logical formula.

Here $\bigoplus_{i<\alpha} \psi_i = \psi_0 + \psi_1 + ... + \phi_n$, where $+$ indicates *exclusive disjunction*. This is a multi-argument truth-function, defined as follows:

$\bigoplus_{i<\alpha} \psi_i$ is true iff *exactly one* of the ψ_i-terms is true, and all the others, false.

Exclusive disjunction it the 'old fashioned' kind of disjunction. It can be found in e.g. the work of Immanuel Kant, in the Table of Judgements from the Critique of Pure Reason, as the third judgement of Relation (M. van Lambalgen, pers. comm.). For purposes of logical deduction it is not a particularly easy-to-handle operation. It is not analysable: $p + q + r$ is not the same as $(p + q) + r$, as can be seen by taking p, q, and r all true. Furthermore, the simple introduction rule for $\vee$ has no equivalent for $+$ (see also Quarfood, 2013)

Definition 3.3. *Inferential sentence*

A formula of the form:

$$\phi \Rightarrow \psi$$

with ϕ a conceptual formula and ψ a logical formula, will be called an *inferential sentence*.

Many inferential sentences that we shall actually use are of the form:

$$\phi \Rightarrow \bigoplus_{i\leq n} \psi_i$$

with every ψ_i a conceptual sentence. In which case we shall speak of an *expansion*.

Definition 3.4. *Realizability sentence*

A formula of the form:

$$\Diamond\phi$$

with ϕ a conceptual formula, will be called a *realizability sentence*.

Definition 3.5. *Antecedental and consequential arguments*

In the inferential sentence:

$$\phi \Rightarrow \psi$$

the arguments in $arg(\phi)$ will be called the *antecedental* arguments, and those in $arg(\psi) - arg(\phi)$, the *consequential* arguments.

Now we can turn to the semantics of Φ. Given an language $\mathcal{L}$ of Φ we can define a model $\mathcal{M} = \langle W, D, I, L\rangle$, where:

- W is the set of possible worlds.
- D is the set of disjoint domains D_w of individuals, one for each world $w \in \mathcal{W}$.
- I is the interpretation-function, assigning a subset $I(P) \subseteq D^n$ to each predicate $P \in \mathcal{P}$, where n is the arity of P.
- L is the set of basic inferences ('laws').

There is also a set $\mathcal{P}_B \subseteq \mathcal{P}$ of *basic* predicates. These are the predicates figuring in the inferential sentences in L.

The accessibility relation is omitted, for we shall simply take each world to be accessible from each world. In accordance with the understanding of modality described above, kinds will run across the whole domain of *possible* individuals. Assignments will take care of keeping the different worlds separate.

Definition 3.6. *Assignment*

An *assignment* $\nu \in (D_w)^{\mathbf{x}}$ is a function from a set of variables to the individuals belonging to one and the same world $w \in W$.

Definition 3.7. *Extension of an assignment*

The assignment ν' is an *extension* of the assignment ν iff:

- $dom(\nu) \subseteq dom(\nu')$.
- for every $x_i \in dom(\nu)$: $\nu(x_i) = \nu'(x_i)$.

Definition 3.8. *Realization of a conceptual formula*

The conceptual formula ϕ is *realized for assignment* ν *by assignment* ν' iff one of the following conditions obtains:

- $\phi = x$ and ν' is an extension of ν, such that $\{x\} \subseteq dom(\nu')$.
- $\phi = R(\mathbf{x})$ with $R \in \mathcal{P}$ and ν' is an extension of ν, such that $\mathbf{x} \subseteq dom(\nu')$, and $\langle \nu'(x_1)...\nu'(x_n) \rangle \in I(R)$.
- $\phi = \psi \cdot \chi$ and both ψ and χ are realized for ν by ν'.

The conceptual formula ϕ is *realized for assignment* ν iff it is realized for ν by any extension ν'.

The idea behind realization is that the assignment fixes individuals from D for some of the arguments, whereas those for the remaining arguments may be chosen freely so as to make the formula come out true. For a conceptual formula the extension must be one and the same throughout all subformulas. For a logical formula this need not be so.

Definition 3.9. *Realization of a logical formula*

The logical formula ϕ is *realized for assignment* ν iff one of the following conditions obtains:

- ϕ is a conceptual formula and ϕ is realized for ν.
- ϕ is not a conceptual formula and:

- $\phi = \neg\psi$ and ψ is not realized for ν.
- $\phi = \psi \cdot \chi$ and both ϕ and χ are realized for ν.
- $\phi = \psi \vee \chi$ and either ϕ or χ are realized for ν.
- $\phi = \bigoplus_{i \leq n} \psi_i$ and for exactly one $i \leq n$: ψ_i is realized for ν.

Definition 3.10. *Validity of an inferential sentence*

The inferential sentence:

$$\phi \Rightarrow \psi$$

is valid iff, for every assignment ν with $dom(\nu) = arg(\phi)$ for which ϕ is realized, ψ is realized.

Definition 3.11. *Validity of a realizability sentence*

The realizability sentence:

$$\Diamond\phi$$

is valid iff there is an assignment ν for which ϕ is realized.

Clearly the sentence $\Diamond\phi$ is valid precisely when $\phi \Rightarrow \bot$ is not.

Finally, the set L, the set of basic 'laws' of the model $\mathcal{M}$, will have as its elements a finite number of so-called *foundational* inferences. Validity of inferences or realizabilities within $\mathcal{M}$ depends all and only on the inferences in this set. The Axioms of Mereology, to be given in Chapter 4, will by default be contained in L.

3.4 Synonymy

If two kinds, P and Q, are necessarily co-extensive, i.e.:

$$P(x) \Leftrightarrow Q(x)$$

can we say that both are *just the same* kind, $P = Q$? First of all, if this were the case, what would it *mean*? For abundant properties, i.e. sets, we have little else than extension as a guideline. The acceptance of kinds often goes together with a realist stance towards the universals in question. If behind kinds there are really entities that manifest their existence in the objects instantiating them, then these universals can take care of themselves, and there seems to be no obvious reason why two *different* universals could not have *the same* extension. But even the nominalist—the one rejecting the reification of universals—may be sympathetic to making the distinction.

Although abundance and extensionality have the same 'feel' about them, it is important to appreciate the difference. Abundance holds that to every extension there is *at least* one property, whereas extensionality says that there is *at most* one such property. Extensionalism, the doctrine that extension is all there is to properties, breeds no interest in the distinction, but precisely the rejection of extensionalism, as advocated here, should make one sensitive to details of this sort. The issue is whether it is possible that:

$$HasHeart(x) \Leftrightarrow HasKidneys(x)$$

and yet:

$$HasHeart(x) \neq HasKidneys(x)$$

This, I think, is a real question. It is conceivable that, if on the basic level—i.e. that of the set L of basic inferences of the model—no co-extensional predicates are introduced, then extensionality follows. It is also conceivable that non-identical co-extensive concepts can be formed under 'favourable' circumstances. A proof for one of these options would be a substantial discovery.

Be all this as it may, it will not stretch our charity too far to agree that the following case is *not* a counterexample to extensionality:

$$P(x) := A(x) \cdot B(x)$$

$$Q(x) := B(x) \cdot A(x)$$

Since we have decided to accept conjunctive kinds, there can be no objection to these definitions of P and Q; and obviously $P(x) \Leftrightarrow Q(x)$. But this, it seems, is simply because $P = Q$. The variation in the definitions for P and Q is an artefact of our notation, hardly more significant than if they would have been printed in a different font. Cases like this are behind the idea of synonymy in Φ.

Synonymy will be expressed as:

$$\phi \doteq \psi$$

where both ϕ and ψ must be *conceptual formulas*. Synonymy will be defined for these, and thereby derivatively for predicates. To fence off obvious synonymies from co-extensionality issues that are really worth a dispute, we shall axiomatize the notion of synonymy.

Before doing so, there is a small technicality worth mentioning. Suppose that $R(x, y) \doteq S(x, y)$, for a non-symmetric R. In such a synonymy-expression we intend the bonds across the equality-sign to be meaningful. Since R is non-symmetric, it is not the case that $R(x, y) \doteq S(y, x)$. But out of context, as a mere symbol, clearly $R(x, y)$ conveys exactly the same concept as $S(y, x)$, since the arguments only contribute their sameness-difference relations. We shall express this as $R(x, y) \sim S(y, x)$. Neither $\doteq$ nor $\sim$ belong to Φ proper. They are part of a meta-language which can be used to make statements *about* Φ-expressions.

The Axioms of Synonymy are the following:

Axiom 1. *Reflexivity*

For every conceptual formula ϕ: $\phi \doteq \phi$

Axiom 2. *Symmetry*

For every pair of conceptual formulas ϕ, ψ: if $\phi \doteq \psi$ then $\psi \doteq \phi$.

Axiom 3. *Commutativity of conjunction*

For every pair of conceptual formulas ϕ, ψ:

$$\phi \cdot \psi \doteq \psi \cdot \phi$$

Axiom 4. *Associativity of conjunction*

For every triple of conceptual formulas ϕ, ψ, χ:

$$(\phi \cdot \psi) \cdot \chi \doteq \phi \cdot (\psi \cdot \chi)$$

Axiom 5. *Substitution of arguments*

For every pair of conceptual formulas ϕ, ψ, if $\phi \doteq \psi$ then:

$$[y/x]\phi \doteq [y/x]\psi$$

Axiom 6. *One-sided substitution of arguments*

For every pair of conceptual formulas ϕ, ψ, if $\phi \doteq \psi$ then:

$$\phi \doteq [y/x]\psi$$

$(y \notin arg(\phi) \cup arg(\psi))$

Axiom 7. *Conjunctivity*

For every P, Q there is an R such that:

$$P(\mathbf{x}) \cdot Q(\mathbf{y}) \doteq R(\mathbf{x}, \mathbf{y})$$

There are two special cases:

- P and Q are identical. Then for $P(\mathbf{x}) \cdot Q(\mathbf{y})$ we get: $P(\mathbf{x}) \cdot P(\mathbf{x}) \doteq P(\mathbf{x})$.
- Q is a null predicate, with $\mathbf{y} \subseteq \mathbf{x}$. Then for $P(\mathbf{x}) \cdot Q(\mathbf{y})$ we get: $P(\mathbf{x}) \cdot \mathbf{y} \doteq P(\mathbf{x})$.

Axiom 8. *Rearrangement of arguments*

For every $P(\mathbf{x})$, for every array $\mathbf{x}'$ with $arg(\mathbf{x}') = arg(\mathbf{x})$, there is a P' such that:

$$P(\mathbf{x}) \doteq P'(\mathbf{x}')$$

We shall call such an $\mathbf{x}'$ a *rearrangement of* $\mathbf{x}$. This axiom allows us to take a lot more advantage of the array-notation for variables: we can (temporarily) line them up in any convenient order. Notice that it also includes cases like: $P(x, y, y) \doteq P'(x, y)$: the repeated occurrence of the same argument under one predicate can be eliminated by switching to a different predicate.

The axioms 7 and 8 introduce an important element into Φ: the element of *definition*. Even though e.g. the kind R is guaranteed to exist, we will have to invent a name for it. This is why we shall customarily write: $R(x, y) :\doteq P(x) \cdot Q(y)$.

For the next axiom we need some definitions:

Definition 3.12. *Bond*

A *bond* is the occurrence of identical variables across different atomic formulas in one expression.

So in:

$$P(x, y, y) \cdot Q(y) \doteq R(x, z)$$

there is one bond on y, between both atomic formulas on the left-hand side, and one on x, across the identity-sign.

Definition 3.13. *Conservative replacement*

A *conservative* replacement of a formula is one by which no bonds are created or destroyed (i.e. the newly introduced formula has exactly the same bonds with the rest of the expression as the formula to be replaced).

Axiom 9. *Replacement of formulas*

Let $\phi \cdot \xi \doteq \chi$ (ϕ, ξ, χ are conceptual formulas, with ξ possibly a null-formula), and let $\phi \doteq \psi$. Then:

$$\psi \cdot \xi \doteq \chi$$

if the replacement of ψ for ϕ is conservative with respect to the formula on the left-hand side.

Axiom 10. *Bond-Sameness*

A sameness-relation is synonymous to a bond between arguments:

$$\phi(x, y, \mathbf{z}) \cdot Same(x, y) \doteq \phi(x, x, \mathbf{z})$$

We shall use the terms *tightening* and *relaxation* for conceptual formulas in which bonds are created or destroyed, respectively. Axiom 10 provides alternative ways to express tightenings and relaxations: by adding or removing $Same$-elements in a conceptual formula.

A *balanced synonymy* is one in which the set of arguments is the same on both sides of the equality-sign.

Theorem 3.1. *Balanced synonymy is transitive: if $\phi \doteq \psi$ and $\psi \doteq \chi$ (with $arg(\phi) = arg(\psi) = arg(\chi)$) then:*

$$\phi \doteq \chi$$

Proof. This is an immediate consequence of Axiom 9 □

Notice that this is not trivial: $R(x, y) \doteq R(v, w)$ and $R(v, w) \doteq R(y, x)$, but it need not be that $R(x, y) \doteq R(y, x)$.

Chapter 4

Parts and Wholes

4.1 Introduction

Suppose, as is probably right, that the world is composed of a large number of elementary particles, and that those particles are all of a finite number of mutually exclusive kinds. Then there is a sense in which the Identification Problem is solved—do we not thereby have an exhaustive way to subsume all the parcels of the world under kinds?

There is a sense in which we do, but it is a poor sense. Surely there are *more* entities worthy of being called so than only the smallest parts of the world: almost everything we think of as an object is *composed* of these smallest parts. Even though Rosen and Dorr, in a remarkable study (2002), have shown that it is possible, philosophically, to cast doubt on the very idea of composition, it is hard to think of something more intuitive, and we shall not pursue this skeptic line of thought here. Yet this basic fact of ontology, no matter how indubitable and familiar, is at the same time an exceptionally *brute* fact. Given two things, there is (at least in certain cases) a third, which is not identical to any of the first two, but an entirely acceptable object all the same. This third object is there, just because the first two are there. On most accounts it is *nothing over and above* the first two—even though, to repeat, it does not *equal* any of both either!

This deep metaphysical enigma is not particularly hard to describe logically. The theory taking care of this, since the famous pioneering work of Leśniewski (Sinisi, 1983), is called (classical) *mereology*. In its most basic form it is a complemented Boolean algebra, normally with the empty set left out (e.g. Herre, 2010). Since, in this thesis, we want to arrive at an account of how entities *other* than atoms instantiate kinds, we will have to make use of mereology. But first let me explain what I believe the role of mereology should be. Its essence, as I understand it, is that it describes an *extra*-logical phenomenon—the part-whole relations between the things the world is composed of—as good as possible, or aims to do so. Therefore theories that describe whatever features of logical or mathematical structures, however analogous to real part-whole relations, do not *on that ground* deserve to be called mereology, any more than that a theory should be called 'mechanics' for any other reason than its being about the interaction of physical objects.

In the metaphysical literature mereology is often granted a far wider domain of application, viz. as one of the main building-blocs to a new foundation of set theory (Goodman & Quine, 1947; Goodman, 1951; Lewis, 1991). The temptation to use the part-whole relation to bestow some substantiality on the beautiful yet chillingly etheric world of sets is quite pardonable, but as far as I can see, quite futile as well. It is not that the enterprise is *logically* illegitimate. Lewis's (1970, 1991) understanding of (nonempty) subsets of classes as their 'parts' is logically impeccable, but raises the question as to exactly what gain is to be had this from re-naming. In what way do subsets profit from being called 'parts' instead? The answer is of course that it is the *innocence* of mereology that Lewis is after. Mereology includes no commitment to things to which there was no commitment before. The story it tells is ontologically sound and free of paradoxes. And indeed, if we could give set theory a foundation of such metaphysical purety, that would be wonderful! If parthood would eventually turn out to be the thing that was behind our intuitions about set-membership all the time, many problems would be solved at once. We would in all likelihood be in a superb position to round up the sources of paradox one by one.

All this *could have* been true if the story about sets was really in some way about parts and wholes. Yet, to give set theory a full mereological foundation—to describe the membership relation fully in terms of the part-whole relation—is trivially impossible. The structure of set theory is unimaginably richer than that of mereology. Lewis, therefore, needs *more* than only mereology to achieve his goal; he needs the *singleton* as a primitive. Now the fact that much of set theory can be derived from Boolean algebra plus the singleton is a remarkable logical result. I feel no urge whatsoever to raise objections against it. But the singleton is quite an assumption—as Lewis would be the first to admit. Armstrong's (1991) proposal to view singletons, (and, by union, classes general) as states of affairs instantiating unit-determining properties, for the credibility that it has, hardly makes it better, for it *also* needs some additional apparatus. Whatever be of it: mereology *only* will never ground set theory.

As far as I can see the lesson to be learned from this is that mereology is not about sets. I am fully aware that this is dangerous talk. Logics that were invented to be *about* this, have often be discovered, equally defensible, to be *about* something totally different. Isomorphisms are among the most celebrated logical discoveries and amazing progress has been made just when logicians *forgot* about *about*. This is a deep truth. But this truth does not include that every logical system is about everything. Some analogies just do not work! I have no doubts that mereology can be bent and stretched to make it look like set theory. I can also tell the story of the Good Samaritan with such a twist as to make it war propaganda. But that is not an adequate treatment of the story. The story of sets is about *classifying*, which is not joining entities to become one big entity, but bringing them *under one concept*. No zoologist is interested in the big entity that is the whole of all mammals; what is important are the criteria for *being one of them*. And the formal story of parts and wholes is not about classifying, but about *composition*. Wholes are interesting, but not so much because of their logical structure. The innocence of wholes is, after all, primarily due to their *lack* thereof, which is responsible for the annoying fact, for whoever uses wholes as sets, that the individuality of the 'elements' get lost once they are fused. Take a very natural whole, a herd of cows, and observe that the parts are cows, but alas, also horns, hoofs, intestines, etc.; all very much unintended elements of the 'set'. There is in fact disappointingly little that can be said, generally, about the parts of a herd of cows.

Seeking family ties between sets and wholes, therefore, I take to be a mistake, and an enterprise that distracts us from the really interesting relations between both. For intriguing they are: wholes do not make kinds, but they do make new entities, that, by their very nature, must in turn *instantiate* kinds—new kinds! They provide an excellent opportunity to test our ideas about similarity and saturation in new environments.

4.2 Mereology in Φ

Here we shall describe a system of standard (classical) mereology, to be used in Φ-expressions. What is new about the system is its symbolic form; its content is in no way novel. As alluded to above, we shall use a Boolean algebra without null set to provide a semantics for the system or mereology to be used. This commits us—at least for practical purposes—to the existence of truly elementary particles, Democritan atoms. Let $\mathcal{A}$ be the set of possible atoms, and $\mathcal{A}_w \subseteq \mathcal{A}$, the subset of these atoms in world $w \in W$, then for every individual in $d \in D_w$, atomic and composite, there is a non-empty subset of $\mathcal{A}_w$ corresponding to it. The most natural way to proceed is to stipulate that every non-empty subset of $\mathcal{A}_w$ in turn determines exactly one individual $d \in D$, which is the formal version of the already-mentioned principle of unrestricted composition:

$$D_w = \{d : comp(d) \in \mathscr{P}(\mathcal{A}_w) - \{\emptyset\}\}$$

Where *comp* is a function assigning to every individual its (non-empty) subset of $\mathcal{A}_w$.

There is, however a big difference between my version of this principle and, e.g. that of Armstrong. Armstrong (1991) argues for unrestricted mereological composition using the words:

> The Sydney Opera House and $\sqrt{-1}$ have their fusion. (p 192)

I do not believe examples like this one are helpful, since what they ask us to swallow are *exceptionally* unnatural wholes. The squareroot of -1, in the ontology that I would advocate, is not a real entity, hence cannot literally enter into relations, in particular, not into mereological ones. Now if someone should come and say that this is not unrestricted mereological composition, I have no objections. 'My' wholes are effectively wholes of naturalistic parts: particles, fragments of time and space, and combinations of them. I believe that this makes for a far more natural understanding of unrestricted mereological composition.

The primitive mereological predicates are Wh (whole), *Dis* (distinct), *Atom*, and *All*. Their semantics is as follows: let I be the interpretation function of the model $\mathcal{M}$, then:

- $\langle d_1, d_2, d_3\rangle \in I(Wh)$ iff $comp(d_1) \cup comp(d_2) = comp(d_3)$.
- $\langle d_1, d_2\rangle \in I(Dis)$ iff $comp(d_1) \cap comp(d_2) = \emptyset$.
- $d \in I(Atom)$ iff $\#comp(d) = 1$.
- $d \in I(All)$ iff $comp(d) = \mathcal{A}_w$ for some $w \in W$.

There is relatively broad consensus that mereological predications are special in that they involve no genuine addition to what is predicated. In David Lewis's words (Lewis, 1991):

> The fusion is nothing over and above the cats that compose it. It just *is* them. They just *are* it. (p 81)
>
> I say that composition—the relation of part to whole, or, better, the many-one relation of many parts to their fusion—is like identity. The 'are' of composition is, so to speak, the plural form of the 'is' of identity. Call this the Thesis of *Composition as Identity* (p 82)

(See also Bøhn, 2011; Sider, 2007) Another way of stating this intuition is to say that mereological predications are *analytic* in the Kantian sense. Extensionality is an example:

the state of affairs $Same(v,w)$ does not merely follow from $Wh(x,y,v) \cdot Wh(x,y,w)$—the fact that both have the same constituents—but is already present *in* the latter. Purely mereological biconditionals between conceptual formulas are more than mutual entailment-relations: they are the same thing! Therefore I shall treat such relations as *synonymies*. Thus from:

$$Wh(x,y,v) \cdot Wh(x,y,w) \Leftrightarrow Wh(x,y,v) \cdot Wh(x,y,w) \cdot Same(v,w)$$

since both sides are purely mereological, it follows that:

$$Wh(x,y,v) \cdot Wh(x,y,w) \doteq Wh(x,y,v) \cdot Wh(x,y,w) \cdot Same(v,w)$$

For the above semantics I shall now give a system of deduction in Φ. In the literature we can find several axiomatic schemes for mereology; here I want to take the concept of whole, rather than part, as my starting-point. The first five axioms, designed to put to work the predicate Wh, cover what is the core of most mereological theories:

Axiom 1. *For any non-empty array of entities* $\mathbf{x}$ *there is a further entity, w, which is the* whole *(Wh) of the entities in* $\mathbf{x}$*:*

$$\mathbf{x} \doteq Wh(\mathbf{x},w)$$

where $Wh(\mathbf{x},w) \doteq Wh(\mathbf{x}',w)$ *if* $\mathbf{x}'$ *is a rearrangement of* $\mathbf{x}$.

This is a peculiar synonymy, first because it equates a bare formula to a predicative one, and second, because w occurs only on the right-hand side. Both should, I think, be thought of as the exclusive 'right' of the predicate Wh and its relatives.

We shall furthermore define $Same(x,y) :\doteq Wh(x,y)$, bringing $Same$ and Wh under one roof.

Axiom 2. *A given array of entities has only one whole:*

$$Wh(\mathbf{x},v) \cdot Wh(\mathbf{x},w) \Rightarrow Same(v,w)$$

Axiom 3. *All entities are wholes:*

$$x \doteq Wh(\mathbf{v},x)$$

Axiom 4. *Any whole is also the whole of all but one of its parts and itself:*

$$Wh(\mathbf{x},y,w) \doteq Wh(\mathbf{x},w,w)$$

Axiom 5. *Wholes add up to bigger wholes*

$$Wh(\mathbf{x},p) \cdot Wh(p,\mathbf{y},w) \doteq Wh(\mathbf{x},\mathbf{y},w)$$

As for the next predicate to be put on stage, $Part$, we shall also treat that as a notational variant of Wh:

$$Part(x,y) :\doteq Wh(x,y,y)$$

Using the above axioms we can be prove:

Lemma 4.1. *Constituents are parts:*

$$Wh(x,\mathbf{y},w) \Rightarrow Part(x,w)$$

Proof. $Wh(x, \mathbf{y}, w)$

$\doteq Wh(\mathbf{y}, x, w)$ Axiom 1

$\doteq Wh(\mathbf{y}, p) \cdot Wh(p, x, w)$ Axiom 5

$(...) \Rightarrow Wh(x, p, w)$ Axiom 1

$(...) \Rightarrow Wh(x, w, w)$ Axiom 4

$(...) \Rightarrow Part(x, w)$ Def. *Part*

Axioms 1 and 4 □

Lemma 4.2. *The whole of parts is itself a part:*

$$Part(x, w) \cdot Part(y, w) \cdot Wh(x, y, v) \Rightarrow Part(v, w)$$

Proof.

$Part(x, w) \cdot Part(y, w) \cdot Wh(x, y, v)$

$\doteq Wh(x, w, w) \cdot Wh(y, w, w) \cdot Wh(x, y, v)$ Def. *Part*

$\doteq Wh(x, w, w) \cdot Wh(w, y, w) \cdot Wh(x, y, v)$ Axiom 1

$\doteq Wh(x, y, w, w) \cdot Wh(x, y, v)$ Axiom 5

$\doteq Wh(x, y, p) \cdot Wh(p, w, w) \cdot Wh(x, y, v)$ Axiom 5

$(...) \Rightarrow Wh(p, w, w) \cdot Same(p, v)$ Axiom 2

$(...) \Rightarrow Wh(v, w, w) \doteq Part(v, w)$ Def. *Part*

□

Theorem 4.1. *The relation $Part(x, y)$ is reflexive, antisymmetric, and transitive.*

Proof.

Reflexivity:

$x \doteq Wh(v, w, x)$ Axiom 3

$\doteq Wh(v, x, x)$ Axiom 4

$(...) \Rightarrow Part(x, x)$ Lemma 4.1

Antisymmetry:

$Part(x, y) \cdot Part(y, x)$

$\doteq Wh(x, y, y) \cdot Wh(y, x, x)$ Def. *Part*

$\doteq Wh(x, y, y) \cdot Wh(x, y, x)$ Axiom 1

$(...) \Rightarrow Same(x, y)$ Axiom 2

Transitivity: By Axioms 5 and 4:

$Part(x, y) \cdot Part(y, z)$

$\doteq Wh(x, y, y) \cdot Wh(y, z, z)$

$\doteq Wh(x, y, z, z)$ Axiom 5

$\doteq Part(x, z)$ Lemma 4.1

□

The Axioms 6 - 9 are those governing the predicate Dis:

Axiom 6. *There is a predicate Dis, with* $Dis(x,y) \doteq Dis(y,x)$*, indicating the being* distinct *of two entities. Two entities are either distinct, or they overlap:*

$$x, y \Rightarrow Dis(x,y) + Ov(x,y)$$

where $Ov(x,y) :\doteq Part(a,x) \cdot Part(a,y)$*.*

Axiom 7. *If something is distinct from two things, it is distinct from their whole:*

$$Dis(a,x) \cdot Dis(a,y) \cdot Wh(x,y,w) \Rightarrow Dis(a,w)$$

Analogous to $Wh(\mathbf{x}, w)$ we shall use multi-argument predicates:

$$Dis(\mathbf{x}) :\doteq \bigodot_{i,j \leq n, i \neq j} Dis(x_i, x_j)$$

In line with this also $Dis(x)$, despite its lack of content, will be used for notational purposes. Furthermore we shall use $WD(\mathbf{x}, w) :\doteq Wh(\mathbf{x}, w) \cdot Dis(\mathbf{x})$.

Axiom 8. *Of two things one is either part of the other, or exceeds it:*

$$x, y \Rightarrow Part(x,y) + Exc(x,y)$$

where $Exc(x,y) :\doteq Part(a,x) \cdot Dis(a,y)$*.*

Axiom 9. *Excess includes complementation:*

$$Exc(x,y) \doteq Wh(x,y,w) \cdot WD(a,y,w)$$

We shall give a couple of useful lemmas, culminating in Theorem 4.2. The proofs can be found in Appendix B.

Lemma 4.3. *If something is distinct from something else, it is distinct from its parts:*

$$Dis(x,y) \cdot Part(a,y) \Rightarrow Dis(a,x)$$

Lemma 4.4. *Parthood plus reverse excess makes proper parthood:*

$$Part(x,y) \cdot Exc(y,x) \doteq PPart(x,y)$$

where $PPart(x,y) :\doteq WD(x,a,y)$*.*

Lemma 4.5. *Parthood implies sameness or proper parthood:*

$$Part(x,y) \Rightarrow Same(x,y) + PPart(x,y)$$

Lemma 4.6. *Complements are unique:*

$$WD(x,a,w) \cdot WD(x,b,w) \Rightarrow Same(a,b)$$

Lemma 4.7. *If something is part of a whole but distinct from one constituent, it is part of the other:*

$$Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x) \Rightarrow Part(a,y)$$

Lemma 4.8. *Mutual excess implies proper overlap or distinctness:*

$$Exc(x, y) \cdot Exc(y, x) \Rightarrow POv(x, y) + Dis(x, y)$$

where $POv(x, y) :\doteq Dis(a, b) \cdot WD(a, u, x) \cdot WD(b, u, y)$ *("Proper Overlap").*

Taking all the pieces together will finally give us:

Theorem 4.2.

$$x, y \Rightarrow Same(x, y) + PPart(x, y) + PPart(y, x) + POv(x, y) + Dis(x, y)$$

The significance of Theorem 4.2 is that it provides a full expansion in terms of the predicates Wh and Dis.

Next we will introduce rules for the remaining primitive predicates $Atom$, and All.

Axiom 10. *There is a predicate Atom:*

$$x \Rightarrow Atom(x) + Complex(x)$$

where $Complex(x) \doteq WD(a, b, x)$

Lemma 4.9. *Two atoms are either identical, or distinct:*

$$Atom(x) \cdot Atom(y) \;\Rightarrow\; Same(x, y) + Dis(x, y)$$

Proof. Immediate from Theorem 4.2 □

Axiom 11. *There is a predicate All, for which:*

$$All(x) \cdot y \Rightarrow Part(y, x)$$

Chapter 5

Composite Kinds

5.1 Introduction

Here is the most simplistic understanding of composite kinds: concepts of them are nothing more than conjunctions of (more) fundamental concepts. Despite its simple-mindedness, let us see how far this idea will lead us. Thus suppose we have $P(x)$ and $Q(y)$. By Axiom 7 of Synonymy, there is a R, such that:

$$R(x,y) :\doteq P(x) \cdot Q(y)$$

R is a genuine relational kind. It may not look like a relation, but the only objection against R is that it is a very uninteresting relation, which, technically, comes down to the—in itself interesting—fact that R is *analysable*: it decomposes (by definition) into separate predications of its relata[1].

There is a sense in which R could be called a composite kind, but it is not this sense that I want to spend the term on. A 'real' composite kinds, I maintain, is something deeper. Describing complex states of affairs by summing up a (large) number of more elementary states of affairs is fine, but the composite kind itself is the *monadic* kind instantiated by the *whole* of this state of affairs, due to all this being the case. So to turn R into a kind like that, the first thing to do is bring this whole on stage. Fortunately our system of mereology supports this move. By Axiom 7 of Synonymy and Axiom 1 of Mereology:

$$P(x) \cdot Q(y) \doteq P(x) \cdot Q(y) \cdot x \cdot y \doteq P(x) \cdot Q(y) \cdot Wh(x,y,w)$$

and again we can bring this under one predicate:

$$S(x,y,w) :\doteq P(x) \cdot Q(y) \cdot Wh(x,y,w)$$

With this we have, as it seems, approached our goal as close as we could with the current apparatus. We have a predication of the whole w to the effect that it is composed of a P and a Q. All we still want is leave out reference to x and y separately. What would be needed to turn the relational $S(x,y,w)$ into a mere *property* of the whole w?

It may be useful to paint the picture in more intuitively appealing hues. Suppose that in fact:

$$S(x,y,w) :\doteq Red(x) \cdot Circle(y) \cdot Wh(x,y,w)$$

[1]Interesting relations, like $Sim_{\mathbf{P}}(x,y)$ are unanalysable: although there are many conjunctions from which they follow (like $P_1(x) \cdot P_1(y)$), they are not itself synonymous to any of them.

Now could there be a predicate, C, such that:

$$C(w) :\doteq Red(x) \cdot Circle(y) \cdot Wh(x, y, w)$$

is a valid definition? Let us not be distracted by the fact that only the argument w occurs at the left-hand side of the synonymy. What it purports to express is that being C is the same as *being the whole of a red entity and a circular one.* Clearly the issue is not about the intelligibility of this condition, but it makes sense to ask if it deserves to be expressed by a monadic concept in Φ, i.e. of a composite kind. One precondition we have formulated for answering this question in the positive is that the suggested kind could be one of a set of alternatives in a saturated model. This does not seem to be the case. Consider:

$$C'(w) :\doteq Blue(x) \cdot Square(y) \cdot Wh(x, y, w)$$

Notice that $C(w)$ and $C'(w)$ are *not* alternatives. A whole of a blue circle and a red square instantiates both!

This appears to deprive us from ways to construe a set of alternatives for concepts like $C(x)$. But then we must decide that C is *not* a legitimate predicate in Φ! The problem indeed stems from the attempt to turn the original $R(x, y)$ into a monadic predicate, applicable to the whole of x and y. For no such difficulty besets the relational R itself; it can easily be thought of as belonging to a set of alternatives:

- $R(x, y) \doteq Red(x) \cdot Circle(y)$
- $R'(x, y) \doteq Red(x) \cdot Square(y)$
- $R''(x, y) \doteq Blue(x) \cdot Circle(y)$
- (etc.)

As long as we keep x and y separate, all we have is a conjunction of alternatives, which makes for a set of alternative conjunctions.

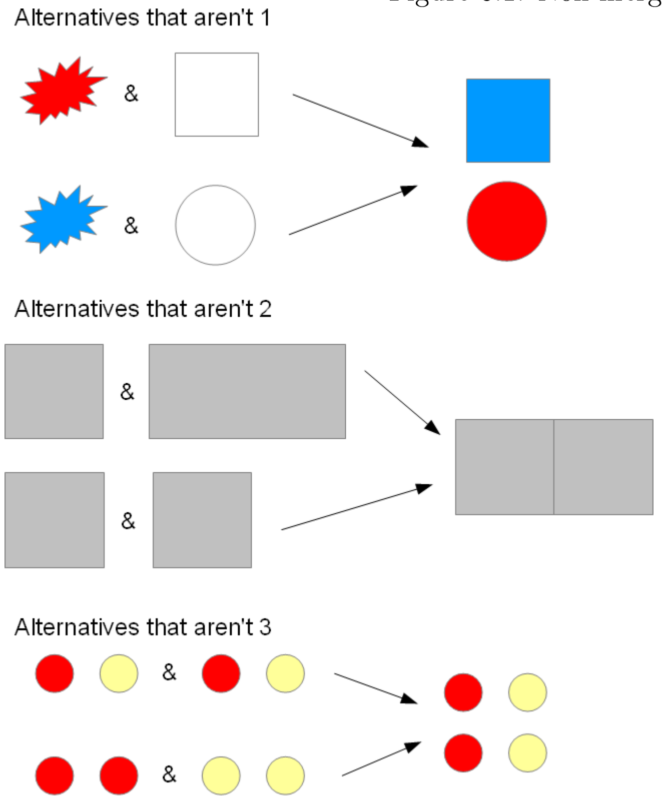
Does this mean that this approach to composite kinds is wrong-headed in principle? Not necessarily; the failure of $Red(x) \cdot Circle(y) \cdot Wh(x, y, w)$ to fit into a similarity system when turned into a monadic predicate of w essentially turns on the fact that *Red* and *Circle* are not of the same *type*. Just try:

$$S(w) :\doteq Red(x) \cdot Yellow(y) \cdot Wh(x, y, w)$$

Here we cannot run the argument as above. This time both x and y are fully determined along the same dimension, namely colour. Being red does not exclude being square, but it does exclude being yellow. Flipping the entities cannot be done.

Let us call a conceptual formula *mergeable* if it can be rephrased as a monadic concept. There are of course more conditions for being a legitimate Φ-concept, like locality (see Chapter 3), but they cause no difficulties; the possibility in principle of saturation is what we will focus on now. The lesson so far to be learned about mergeability seems to be the following: complex concepts have a chance to be mergeable, but *only if* all the predicates in the composition belong to the same type. Every part must be given under a description of the same specificity. All in terms of colour, *or* all in terms of shape, etc.

Figure 5.1: Non-mergeability



5.2 Quasi-atoms

But this will not do yet. What we want is a criterion for how to create composite predicates in such a way that the constituting concepts are *guaranteed* to be mergeable. What we had so far was this: let (**P**) be a type, then alternatives are being created by taking the instances of:

$$P_i(x) \cdot P_j(y) \cdot Wh(x, y, w)$$

($P_i, P_j \in \mathbf{P}$). Under this assumption, however, consider the following concepts of type (**S**) (shapes):

$$S(w) :\doteq Square(x) \cdot Square(y) \cdot Wh(x, y, w)$$

and:

$$S'(w) :\doteq Square(x) \cdot Rect(y) \cdot Wh(x, y, w)$$

(where *Rect* is a whole of two distinct squares). No doubt $Square(x)$ and $Rect(x)$ are alternatives in **S**, but an object consisting of two squares is *also* an object consisting of one square and a 'rect', simply because the first is one half of the second. In the latter case both constituents of the whole *overlap*, but overlapping parts can also make wholes.

The possibility of overlap creates a lot of noise in the system. For the next attempt, therefore, we will demand that there be no. It can reasonably be assumed that there is no loss of generality involved in this: every mereological condition should eventually boil down to a condition of this sort, by omitting the overlapping parts. So this is the new hypothesis: a conceptual formula is mergeable if it is of the form:

$$P_i(x) \cdot P_j(y) \cdot WD(x, y, w)$$

Alas, even this is not enough. Consider pairs of coloured atoms: $RG(x)$ means that x consists of a red, and a green atom. Clearly if $RG(x)$ is part of a set of alternatives, then $RR(x)$ (both red) and $GG(x)$ (both green) are in the same set. Now take:

$$S(w) :\doteq RG(x) \cdot RG(y) \cdot WD(x, y, w)$$

and:

$$S'(w) :\doteq RR(x) \cdot GG(y) \cdot WD(x, y, w)$$

But being an S does not exclude being an S'; quite the contrary: they are the same!

It begins to look like we will have to go down to the very atoms and their properties and relations: the most specific analysis. Now *if* we embrace the assumption of atomicity we have every reason to do so, but this may be yielding to that assumption too soon. And in fact the news is a little bit better: instead of proper atoms (having no proper parts) every *quasi-atomic* kind will also do as a basic unit.

Definition 5.1. *Quasi-atomic kind*

A quasi-atomic kind is a monadic kind P such that the following inference is valid:

$$P(x) \cdot P(y) \Rightarrow Same(x, y) + Dis(x, y)$$

Quasi-atomic kinds are kinds that share with proper atoms the feature that they do not overlap. Being a marble is quasi-atomic, and being a person is. Real (physical) atoms are not proper atoms, but they are quite tolerable quasi-atoms. And proper atoms are quasi-atoms, of course!

We shall now prove that quasi-atoms can indeed serve as units for composite kinds. This proof will require a few intermediate steps.

Definition 5.2. *Multitude of a concept*

A concept $\mu(\mathbf{x})$ is a *multitude* of a concept $P(x)$ iff, for some $n \in \mathbb{N}$:

$$\mu(\mathbf{x}) = \bigodot_{i \leq n} P(x_i)$$

$(n = \#\mathbf{x})$

Definition 5.3. *Specific multitude of a concept*

A concept $\mu(\mathbf{x})$ is a *specific* multitude of a concept $P(x)$ iff, for some $n \in \mathbb{N}$:

$$\mu(\mathbf{x}) = \bigodot_{i \leq n} P(x_i) \cdot \bigodot_{i,j \leq n; i \neq j} Dis(x_i, x_j)$$

$(n = \#\mathbf{x})$

Theorem 5.1. *Let $\mu(\mathbf{x})$ be a specific multitude of a quasi-atomic $A(x)$. Then any A that is part of the multitude equals exactly one of its explicit A-parts:*

$$\mu(\mathbf{x}) \cdot A(y) \cdot Wh(\mathbf{x}, w) \cdot Part(y, w) \Rightarrow \bigoplus_{i \leq n} Same(x_i, y)$$

I.e. there are no A-parts inside the multitude but those referred to by the subformulas. Of course specific multitudes were especially devised to have this property.

Proof. By induction on the number of occurrences of $A(...)$ in $\mu(\mathbf{x})$.

- Base case: $A(x) \cdot A(y) \cdot Wh(x, w) \cdot Part(y, w)$

  $\doteq A(x) \cdot A(y) \cdot Same(x, w) \cdot Part(y, w)$

  $\doteq A(x) \cdot A(y) \cdot Part(y, x)$

  By quasi-atomicity:

  $A(x) \cdot A(y) \cdot Part(y, x)$

  $\Rightarrow (Same(x, y) + Dis(x, y)) \cdot Part(y, x)$

  $(...) \Rightarrow Same(x, y) \cdot Part(y, x) + Dis(x, y) \cdot Part(y, x)$

  As the second disjunct is contradictory:

  $(...) \Rightarrow Same(x, y)$

- Inductive step: Assume that, for n occurrences of $A(...)$ in $\mu(\mathbf{x})$:

$$\mu(\mathbf{x}) \cdot A(y) \cdot Wh(\mathbf{x}, w) \cdot Part(y, w) \Rightarrow \bigoplus_{i \leq n} Same(x_i, y)$$

  Then, by quasi-atomicity:

  $\mu(\mathbf{x}) \cdot A(x_{n+1}) \cdot A(y) \cdot WD(\mathbf{x}, x_{n+1}, w) \cdot Part(y, w)$

  $\Rightarrow \mu(\mathbf{x}) \cdot A(x_{n+1}) \cdot A(y) \cdot Same(x_{n+1}, y) \cdot WD(\mathbf{x}, x_{n+1}, w) \cdot Part(y, w)$

  $+\mu(\mathbf{x}) \cdot A(x_{n+1}) \cdot A(y) \cdot Dis(x_{n+1}, y) \cdot WD(\mathbf{x}, x_{n+1}, w) \cdot Part(y, w)$

  with Lemma 4.7 with respect to the second disjunct:

  $(...) \Rightarrow Same(x_{n+1}, y) \cdot Dis(\mathbf{x}, y)$

  $+\mu(\mathbf{x}) \cdot A(y) \cdot Wh(\mathbf{x}, v) \cdot Part(y, v) \cdot Dis(x_{n+1}, y)$

  and with the inductive hypothesis applied to the second disjunct:

  $(...) \Rightarrow Same(x_{n+1}, y) \cdot Dis(\mathbf{x}, y) + \bigoplus_{i \leq n} Same(x_i, y) \cdot Dis(x_{n+1}, y)$

  $(...) \Rightarrow \bigoplus_{i \leq n+1} Same(x_i, y)$

□

For specific multitudes of quasi-atomic concepts it makes sense to define $\#\mu =$ the number of occurrences of $A(...)$, which will be called the *size* of μ.

The following lemma states that specific multitudes of A's can only instantiate one and the same whole if they are of the same size. If so, there must be an *identification* between the arguments of both. Intuitively this is very unsurprising result, but it is most practical for further use.

Lemma 5.1. *Let $\mu(\mathbf{x})$ and $\nu(\mathbf{y})$ both be specific multitudes of some quasi-atomic $A(x)$. Then, if $\#\mu = \#\nu = n$:*

$$\mu(\mathbf{x}) \cdot \nu(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w) \Rightarrow \bigoplus_{j \leq n!} \bigodot_{i \leq n} Same(x_i, y_{f_j(i)})$$

(f_j is the j^{th} isomorphism in $\{1...n\}^{\{1...n\}}$), otherwise:

$$\mu(\mathbf{x}) \cdot \nu(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w) \Rightarrow \bot$$

Proof. As a generalization of Theorem 5.1 we have:

$$\mu(\mathbf{x}) \cdot \nu(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w) \Rightarrow \bigodot_{j \leq n} \bigoplus_{i \leq m} Same(x_i, y_j)$$

$(\#\mu = m,\ \#\nu = n)$. This can be rewritten as:

$$\mu(\mathbf{x}) \cdot \nu(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w) \Rightarrow \bigoplus_{j \leq m^n} \bigodot_{i \leq n} Same(x_i, y_{f_j(i)})$$

(f_j is the j^{th} function in $\{1...m\}^{\{1...n\}}$). Consider the disjuncts in the consequent. They all consist of an identification of every element of $\mathbf{x}$ with one of the elements of $\mathbf{y}$. The identification is done by the function f_j. If this function is not 1 - 1, a contradiction follows from the fact that, by the definition of specific multitude, all elements of $\mathbf{x}$ are distinct. If it is not onto, a contradiction follows from the fact that the whole of $\mathbf{y}$ does not exceed the whole of $\mathbf{x}$. Hence, the only disjuncts that are not contradictory are those for which f_j is an isomorphism, $m = n$, and $\#\mu = \#\nu$. □

Lemma 5.2. *Let $\mu(\mathbf{x})$ and $\nu(\mathbf{y})$ both be specific multitudes of some quasi-atomic $A(x)$. If:*

$$\Diamond\mu(\mathbf{x}) \cdot \nu(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w)$$

then $\mu(\mathbf{x}) \doteq \nu(\mathbf{y})$

Proof. This is an immediate consequence of Lemma 5.1 plus Definition 5.3: specific multitudes of equal size must be synonymous. □

This means that wholes of specific multitudes of quasi-atomic concepts pass the test of allowing saturation. Multitudes of different size are alternatives to one another. Any whole can instantiate only one of them.

It should be noted, however, that this putative similarity system does not run through *all* the individuals in the domain. It only applies to whatever is a multitude of $A(x)$ in the first place. Even though we always have the fall-back assumption of A standing for *proper* atomicity, it is worthwhile to see what we would need to 'save' saturation without it.

To secure saturation, we need two things. First, we need a concept that expresses 'multitude of $A(x)$' in general. Suppose, for the moment, that we have laid hand on such a concept, and it is $A^{\#}(x)$. Then we can write:

$$A^{\#}(x) \Rightarrow A(x) + A^2(x) + A^3(x) + ...$$

where $A^n(x) :\doteq \bigodot_{i \leq n} A(v_i) \cdot WD(\mathbf{v}, x)$. Second, we must ascertain ourselves that the (abstract!) predicate $A^{\#}$ is *itself* part of a similarity system such that:

$$x \Rightarrow A^{\#}(x) + ... + ...$$

In the last section I shall add some ideas about what sort of concept $A^{\#}(x)$ could be. As for the second point: what would be called for is a theorem to the effect that everything whatsoever falls under a certain higher level type of multitude. However interesting, providing this is beyond the ambitions of this thesis.

Putting our trust in the availability of solutions to both issues just mentioned, let us continue the argument about mergeability.

Definition 5.4. *Predicated specific multitude*

Let $\mu(\mathbf{x})$ be a specific multitude of a quasi-atomic concept $A(x)$ and let $\pi(\mathbf{x})$ be any conjunction of predications on subarrays of $\mathbf{x}$ whatsoever, then:

$$\phi(\mathbf{x}) \doteq \mu(\mathbf{x}) \cdot \pi(\mathbf{x})$$

will be called a *predicated specific multitude.* Its *size* $\#\phi$ is the same number as $\#\mu$.

As a special case of predicated specific multitudes, consider a similarity system $\mathfrak{S} = \langle \mathbf{P}, Sim_{\mathbf{P}} \rangle$ of monadic predicates, such that, for some quasi-atomic A, for any $P_i \in \mathbf{P}$, $P_i(x) \doteq A(x) \cdot P'_i(x)$. Then we can define:

Definition 5.5. *Full concept of type* $(\mathbf{P})$ *(monadic case)*

Let $\mathfrak{S}$ be the similarity system just mentioned and let $\mu(\mathbf{x})$ be a specific multitude of a quasi-atomic concept $A(x)$. Furthermore $f \in \{1...\#\mathbf{P}\}^{\{1...\#\mathbf{x}\}}$ is a function assigning alternative predications $P_i \in \mathbf{P}$ to the arguments in $\mathbf{x}$. Then:

$$\phi(\mathbf{x}) \doteq \mu(\mathbf{x}) \cdot \bigodot_{i \leq n} P_{f(i)}(x_i)$$

will be called a *full concept* of type $(\mathbf{P})$. Its *size* $\#\phi$ is the same number as $\#\mu$.

In such concepts the P_i-predications add flesh to the $\mu(\mathbf{x})$-bones. Concepts of this sort show a greater variation than only difference of size. For monadic P_i's it is evident that all non-synonymous concepts are alternatives to one another: since every argument in $\mathbf{x}$ is predicated according to one of the alternatives in $\mathbf{P}$, counting how many of them are P_1, P_2, etc., is all that can be said about their whole.

This is the sort of concept that we shall assume to be mergeable. Of course this is an additional theoretical assumption, in effect a new Axiom of Synonymy, to be introduced below. Just like in the case of Axiom 7, the resulting kind will receive its name by a definition:

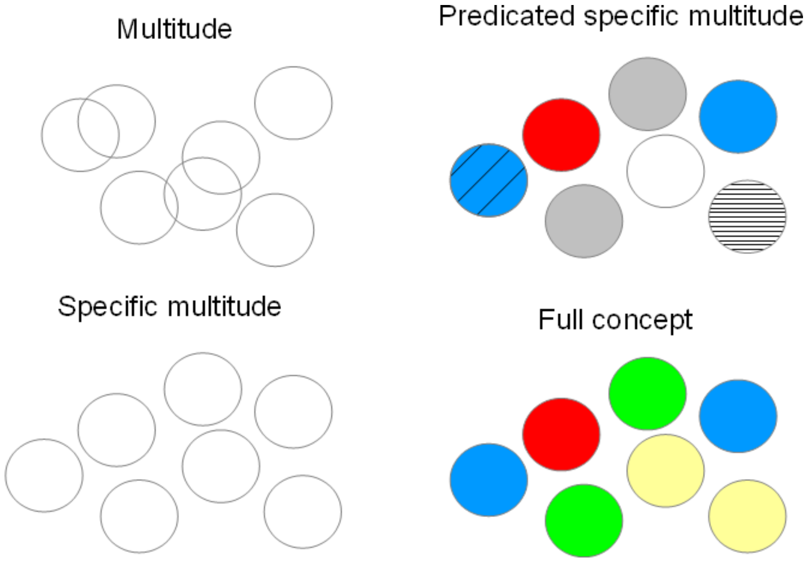
$$S(w) :\doteq \mu(\mathbf{x}) \cdot \bigodot_{i \leq n} P_{f(i)}(x_i) \cdot Wh(\mathbf{x}, w)$$

As long as we have only full concepts based on monadic predicates, however, this new piece of theory is of limited significance. What is missing out is something to seal the quasi-atoms together and make the whole a *real* whole, rather than a relatively arbitrary pile of distinct individuals. For this we need primitive relations.

5.3 Primitive Relations

So far we have considered only primitive properties and supervenient relations, but *primitive relations*, relations that do not follow from any more basic condition, are indispensable for a richer understanding of composite kinds. Primitive relations are marked by the fact that intrinsically identical things can be related by them in different ways. When it comes to assessing the actual abundance of such relations in the real world, one of the remarkable observations is that their number appears to be limited. Hume thought that *time*, *space*, and *causality* were the only cases. Of these, time and space are most susceptible of being modelled in elementary Φ-terms, as long as we contend ourselves with describing the 'manifest image', and even that in simplified and idealized form. We do good to acknowledge some variation in those relations: apart from being merely adjacent to one another, things can be *connected* in a variety of ways. A celebrated example of this is *chemical*

Figure 5.2: From multitude to full concept



bonds, as they occur in e.g. methane. Methane often figures in discussions about structural universals (Lewis, 1986a; Armstrong, 1986; Pagès, 2002; Kalhat, Mormann, 2010, 2012); here we obviously want to be able to account for the *kind* methane as a composite kind.

What we shall prove in this section is that what has been shown to be true for full concepts of monadic types, viz. that they are eligible for being declared mergeable, is true for full concepts generally. First of all our notions of saturation are in need of an update. Saturation for relations is significantly more complex than it is for monadic predicates. We will be concerned with primitive relations that apply to some quasi-atomic kind, let us call it A, and we shall word these relations in terms of predicated specific multitudes. To keep the formal notation from running wild, we shall make use of a convention and introduce a shorthand. The convention is that the predicates of main interest will be those that already include the A-ness of their instances. So[2] $P(x) \doteq P'(x) \cdot A(x)$. Furthermore we shall use 'hats' to indicate relations of being distinct. Consequently, instead of writing:

$$P_3'(x) \cdot A(x) \cdot P'_1(y) \cdot A(y) \cdot Dis(x, y)$$

we shall write:

$$P_3(\hat{x}) \cdot P_1(\hat{y})$$

Remember the monadic case. The simplest of monadic foundational expansions just look like a display of the alternatives in the similarity-system, e.g.:

(22) $A(x) \Rightarrow P_1(x) + P_2(x) + P_3(x)$

We could also have written:

(23) $A(x) \Rightarrow \bigoplus_{P \in \mathbf{P}} P(x)$

For relational systems, however, the simple parallelism between type and expansion breaks down. As an illustration take the following system. Its main concept is a dyadic primitive relation:

$$R(x, y)$$

[2]An even shorter way to express this is: $P(x) \doteq P(x) \cdot A(x)$

that can be used as a primordial element in a model for (one-dimensional) spacial or temporal relations. R, of type ($\mathbf{R}$), allows the formation of *chains* of different length. Apart from being somewhere inside a chain, an entity can be a left or right *terminus*. First we have:

(24) $A(x) \Rightarrow T_r(x) + R(\hat{x}, \hat{y})$

Given an A, either it is a right terminus, or there is a next A in the chain, to the right. Notice that the implicit quantification of the consequential argument y is existential: *there is* a next A. Now (24) cannot be the whole story about this type, since there is no mention of the antecedental argument x occurring in the *second* position of R. So, additionally:

(25) $A(x) \Rightarrow T_l(x) + R(\hat{v}, \hat{x})$

The expansion displaying *all* the alternatives of type ($\mathbf{R}$) should be one comprising *all* possible alternatives with respect to *every* predicate of this type. In this case it can be calculated from (24) and (25) by working out the conjunction of the consequents of both, to arrive at:

(26) $A(x) \Rightarrow T_r(x) \cdot T_l(x) + T_r(\hat{x}) \cdot L(\hat{x}, \hat{y}) + T_l(\hat{x}) \cdot L(\hat{v}, \hat{x}) + L(\hat{v}, \hat{x}) \cdot L(\hat{x}, \hat{y})$

To state the set of alternatives we have to turn all four concepts into single predicates:

- $R_1(x) :\doteq T_r(x) \cdot T_l(x)$
- $R_2(x, y) :\doteq T_r(\hat{x}) \cdot R(\hat{x}, \hat{y})$
- $R_3(x, v) :\doteq T_l(\hat{x}) \cdot R(\hat{v}, \hat{x})$
- $R_4(x, v, y) :\doteq R(\hat{v}, \hat{x}) \cdot R(\hat{x}, \hat{y})$

Two important things should be noticed now. First, in order to retain the information which of the arguments is *antecedental*, we had to choose these definitions so as to put it first: x appears on the first position in all the $R_i(...)$'s. This is mandatory: the specificity of the alternatives in (26) concerns the antecedental argument: *only* with respect to x they represent all the options. This is different for the consequential arguments: these are not split out to full specificity.

Second, it would be unpractical to take this set of alternatives to be the set defining the type. The predicate R itself, which is obviously of type ($\mathbf{R}$), is not among the alternatives. Therefore we shall take $\mathbf{R} = \{R, T_r, T_l\}$ instead: the set of *basic predicates* of the type. Unlike in the typical monadic case, $\mathbf{R}$ cannot be used as the set of alternatives. The set of alternatives will now be called $\mathfrak{A}(\mathbf{R}) = \{R_1, R_2, R_3, R_4\}$.

The following two definitions are updated versions of the earlier ones:

Definition 5.6. *Concept of type* ($\mathbf{P}$)

A *concept of type* ($\mathbf{P}$) is any conceptual sentence $\phi(\mathbf{x})$ such that one of the following conditions hold:

- $\phi(\mathbf{x}) \doteq P(\mathbf{x})$, with $P \in \mathbf{P}$.
- $\phi(\mathbf{x}) \doteq \psi(\mathbf{x}) \cdot \chi(\mathbf{x})$, with both $\psi(\mathbf{x})$ and $\chi(\mathbf{x})$ concepts of type ($\mathbf{P}$).

Definition 5.7. *Full concept of type* ($\mathbf{P}$) *(general case)*

Let $\mu(\mathbf{x})$ be a specific multitude of a quasi-atomic concept $A(x)$, and let $f \in \{1...\#\mathfrak{A}(\mathbf{P})\}^{\{1...\#\mathbf{x}\}}$ a function assigning alternative predications to the arguments. Then:

$$\phi(\mathbf{x}) \doteq \mu(\mathbf{x}) \cdot \bigodot_{i \leq n} P_{f(i)}(x_i, \mathbf{y}_i)$$

(for every i: $\mathbf{y}_i \subseteq \mathbf{x}$) will be called a *full concept* of type $(\mathbf{P})$. Its *size* $\#\phi$ is the same number as $\#\mu$.

The crucial point is that, in a full concept, *every* A is predicated according to one of the alternatives in $\mathfrak{A}(\mathbf{P})$, i.e. to full specificity. The result we are eventually after is that this amounts to the concept's being fully specific *as a whole* also.

Although (26) provides all the alternatives with respect to x, it is still in some way incomplete. In expansions of monadic types like (23) every term in the expansion is straightforwardly a full concept of the relevant type, but in (26) this is not so: of the four alternatives only the first is. Now we could expand every term further, to full specificity of v and y also. This would bring only temporary relief, however, since new consequential arguments would appear. An expansion of $A(x)$ into *only* full concepts of type $(\mathbf{R})$ has infinitely many terms. It consists of chains of every length, with x on every possible position in the chain:

$$A(x) \Rightarrow Ch_1^1(x) + Ch_1^2(\hat{x}, \hat{\mathbf{y}}) + Ch_2^2(\hat{x}, \hat{\mathbf{z}}) + ...$$

($Ch_j^i(\hat{x}, \hat{\mathbf{y}})$ is a chain of length i, with x on position j and y, the rest of the chain). This *type-expansion* of type $(\mathbf{R})$ produces yet another set of alternative: $\mathfrak{T}(\mathbf{R}) = \{Ch_1^1, Ch_1^2, Ch_2^2, ...\}$.

So instead of the straightforward expansions of monadic types, in the polyadic case we have several sorts of expansions that are of interest. I leave it here as an open question which of them underlies the 'true' notion of *similarity*. According to the line of thought developed here there must be at least one similarity concept governing the type $(\mathbf{R})$, but there could well be more of them: maybe one for every expansion! Fortunately (?) the truth in this matter does not seem to have many repercussions for what follows.

A practical way to proceed to our proof is now to consider expansions with *more than one* antecedental argument. They can be created in the following way. Here is the basic expansion:

$$A(x) \Rightarrow \bigoplus_{R \in \mathfrak{A}(\mathbf{R})} R(x, \mathbf{v})$$

So with two antecedental arguments we get:

$$A(x) \cdot A(y) \Rightarrow \left(\bigoplus_{R \in \mathfrak{A}(\mathbf{R})} R(x, \mathbf{v}) \right) \cdot \left(\bigoplus_{R \in \mathfrak{A}(\mathbf{R})} R(y, \mathbf{w}) \right)$$

which is:

$$A(x) \cdot A(y) \Rightarrow \bigoplus_{R \in \mathfrak{A}(\mathbf{R}), R' \in \mathfrak{A}(\mathbf{R})} \big(R(x, \mathbf{v}) \cdot R'(y, \mathbf{w}) \big)$$

But this is not quite enough. The terms of this disjunction are not predicated specific multitudes, as they are not specific at all. To fulfil this condition in the new conjunctions we must add, for any pair of arguments a, b, across the conjuncts, the expansion:

$$(Same(a, b) + Dis(a, b))$$

We can do this because it follows from the quasi-atomicity of A. So finally we get:

$$A(x) \cdot A(y) \Rightarrow \bigoplus_{R \in \mathfrak{A}(\mathbf{R}), R' \in \mathfrak{A}(\mathbf{R})} \left(R(x, \mathbf{v}) \cdot R'(y, \mathbf{w}) \cdot \bigodot_{a \in \{x\} \cup \mathbf{v}, b \in \{y\} \cup \mathbf{w}} (Same(a, b) + Dis(a, b)) \right)$$

The set of alternatives will be called $\mathfrak{A}^2(\mathbf{R})$. Clearly we can do this for any number of antecedental arguments.

Now we are ready for matters of real importance. If full concepts of type (P), monadic or otherwise, apply to the same whole, they are synonymous:

Theorem 5.2. *Let $P(\mathbf{x})$ and $Q(\mathbf{y})$ both be full concepts of type $(\mathbf{P})$. If:*

$$\Diamond P(\mathbf{x}) \cdot Q(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w)$$

then $P(\mathbf{x}) \doteq Q(\mathbf{y})$.

Proof. By Lemma 5.1 we know that $\#P = \#Q = n$. Also we have:

$$P(\mathbf{x}) \cdot Q(\mathbf{y}) \cdot Wh(\mathbf{x}, w) \cdot Wh(\mathbf{y}, w) \Rightarrow \bigoplus_{j \leq n!} \bigodot_{i \leq n} Same(x_i, y_{f_j(i)})$$

By hypothesis at least one of the disjuncts in the consequent is realizable; let us say the k^{th} disjunct is. By an isomorphism f_k from $\mathbf{x}$ to $\mathbf{y}$ every element of $\mathbf{x}$ is identified with one of the elements of $\mathbf{y}$. Hence, by repeated application of Axiom 10 of Synonymy there is a rearrangement $\mathbf{x}'$ of $\mathbf{x}$ such that $Q(\mathbf{x}')$. By Axiom 8 there is also a Q' such that $Q(\mathbf{x}') \doteq Q'(\mathbf{x})$. Clearly $Q'(\mathbf{x})$ is a full concept of type $(\mathbf{P})$, so $Q' \in \mathfrak{A}^n(\mathbf{P})$. But then, as $\mathbf{x}$ can only instantiate one alternative of $\mathfrak{A}^n(\mathbf{P})$, it must be that $P(\mathbf{x}) \doteq Q'(\mathbf{x})$. With $Q(\mathbf{x}') \doteq Q'(\mathbf{x})$ and Theorem 3.1 we have $P(\mathbf{x}) \doteq Q(\mathbf{x}')$. Now of $\mathbf{y}$ we know of its i^{th} argument that either $y_i = x'_i$, or $y_i = z$ with $z \notin \mathbf{x}'$; otherwise a contradiction follows. Therefore, with Axiom 6 of Synonymy, $P(\mathbf{x}) \doteq Q(\mathbf{y})$ □

This removes the final obstacle to declare full concepts of some type to be mergeable: to become monadic concepts of the same type. Notice that, just like the whole was nothing over and above its parts, the monadic concept of the whole does not add anything that was not there before. Therefore it is apt to treat the equivalence between both as a *synonymy.* But it is a new sort of synonymy, not yet in the list of Axioms of Synonymy in Chapter 3. Therefore we need an additional axiom:

Axiom 11. *Let $\phi(\mathbf{x})$ be a full concept of type $(\mathbf{P})$, then there is a monadic concept C such that:*

$$C(w) \doteq \phi(\mathbf{x}) \cdot Wh(\mathbf{x}, w)$$

The last thing that is still lacking is the *similarity system* unifying this new set of predicates. What we have now is, for every number n, a set of alternative monadic predicates for merged wholes of size n. This fixing of n, however, does not look too natural. A far more elegant similarity system would be one comprising *all* the merged versions of full concepts of type $(\mathbf{P})$, for they are all alternatives. This set will be called $\mathfrak{W}(\mathbf{P})$. Thus we have the similarity system $\mathfrak{S} = \langle \mathfrak{W}(\mathbf{P}), Sim_{\mathfrak{W}(\mathbf{P})} \rangle$.

Still one little question: what is the concept that is expanded in this way? We have:

$$C(x) \Rightarrow \bigoplus_{P \in \mathfrak{W}(\mathbf{P})} P(x)$$

What does C stand for? It must be *being a number of A's.* But we have not yet ways to create such a predicate! Intuitively it appears to be an option that is called for. In the last section we will have a look at this issue.

5.4 Numerical Abstraction

This is the idea developed in the last section: suppose that $P_1(\hat{x}) \cdot P_2(\hat{y})$ is a full concept of type (**P**), then there is a predicate S (also of type (**P**)) such that $S(w)$ applies to the whole w of x and y. But what sort of predicate is S? If we have:

$$S(w) \doteq Red(\hat{x}) \cdot Yellow(\hat{y}) \cdot Wh(x, y, w)$$

then $S(w)$ means that w is neither red, nor yellow, but has a colour-state based on its being composed of a red and a yellow part. But now consider:

$$S'(w) \doteq Red(\hat{x}) \cdot Red(\hat{y}) \cdot Wh(x, y, w)$$

This time *both* constituents are red. Now would it be correct to call w red as well; i.e. could we embrace the following inference:

$$S'(w) \Rightarrow Red(w)$$

Intuition will be inclined to accept this—join red parts and there will be a red whole—but we must be very cautious. So far there is nothing that *logically* commits us to the transferability of predicates in this way. Nothing in fact logically commits us to accept that composite kinds need *ever* be the same as those occurring in the composition.

This is the problem before us: there are predicates, like *Red*, that we use for certain entities. Now we would like to be able to re-use the same predicates for the wholes that the said entities are composed of. This seems to be an entirely reasonable desire: it would be odd if every layer of mereological composition would have its own stock of predicates. If something can be *Red*, then, if it is part of a bigger whole, there is at least the conceptual possibility that that whole be every bit as *Red*.

The latter remark, though making good rhetoric, deserves to be qualified a bit. If we take the predicate *Circle* instead of *Red* the argument looses much of its swing. Whatever has only red parts is arguably red, but whatever has only circular parts need not be circular at all. Yet for shapes a very similar point is in order: wholes have shapes that strongly depend on the shapes of their parts. What we are looking for—and is by no means trivial—is an understanding of *compositional abstraction*, i.e. of the way in which properties of a certain type necessitate the re-appearance of properties of that type on higher mereological levels. Notice that in non-mereological models this very idea has no force: there are just entities and some of them instantiate, say, property P, by mere stipulation. However, once we have wholes we have entities that are new, but not entirely novel. Metaphysically they are nothing over and above their parts, so information about those parts *is* information about the whole. Hence the re-appearance of P in wholes is something that, in certain cases, *we should have anticipated*. But how do we anticipate things like this?

Let us see what options we have. Consider:

- $S(w) \doteq Red(\hat{x}) \cdot Yellow(\hat{y}) \cdot Wh(x, y, w)$
- $S'(w) \doteq Red(\hat{x}) \cdot Red(\hat{y}) \cdot Wh(x, y, w)$
- $S''(w) \doteq Yellow(\hat{x}) \cdot Yellow(\hat{y}) \cdot Wh(x, y, w)$

In his *Aufbau*, Carnap endorsed a system in which classes of wholes were arranged according to their having at least one part of the relevant colour. This would place the items falling under S in the same colour-class as those falling under S', but it would also unite the items falling under S and S''. The advantage of this system is that it gives us a simple and straightforward rule for compositional abstraction. The bad news is that, as is easy to see, it instantly destroys saturation: all colour classes start to blend once the wholes get bigger. Whatever has red parts, is *Red* in this sense. We therefore loose the ability to distinguish between objects, like the Red Flag and the Spanish Flag, that are distinguishable beyond any reasonable dispute. The more natural view, as I see it, is that $S'(w)$ entails that w is in fact red, $S''(w)$, that it is yellow and that S is simply a new alternative. This view poses no threat to saturation, but it does rise new issues. What will the *general* rule look like?

Actually to lay down such a general rule is a fairly ambitious project, but let us sketch some contours. A compositional abstraction is, first of all, an abstraction. It is a predicate P_{abs} unifying number (infinitely many, in this case) of underlying predications:

$$\begin{cases} P_{abs}(x) \Rightarrow \bigoplus_{i\in\mathbb{N}} P_i(x) \\ P_i(x) \Rightarrow P_{abs} & (i \in \mathbb{N}) \end{cases}$$

An obvious intuition behind compositional abstraction is that it is closed under mereological composition. The whole of something P_i and something P_j will again instantiate P_k (for some k). Or, in terms of the abstraction itself:

$$P_{abs}(x) \cdot P_{abs}(y) \cdot Wh(x, y, w) \Rightarrow P_{abs}(w)$$

The easy road forward would be to declare certain predicates to behave like this, constituting the class of 'compositional predicates' or something of the kind, but there are good reasons to take the analysis a bit further. First, a stipulation like this would call into being a class of abstractions, among which *Red*, with no concrete predicates underneath; and although there is no legal code forbidding this, it is definitely against the general spirit we have so far reasoned by. Second, the behaviour of, e.g. shapes is different, but nevertheless related, and it would be nice if this could somehow be mirrored in the theory.

Here is how we could take the analysis one step deeper. If the predicate *Red* is an abstraction, then, according to our present line of reasoning, there must be a set of predicates, say $\{Red^*_1, Red^*_2, Red^*_3, ...\}$ of which it is an abstraction. A promising interpretation of this is that there are quasi-atoms A that can be Red^*, which predicate exclusively applies to A's. Furthermore:

- $Red^*_1(w) \doteq Red^*(w)$
- $Red^*_2(w) \doteq Red^*(\hat{x}) \cdot Red^*(\hat{y}) \cdot Wh(x, y, w)$
- $Red^*_3(w) \doteq Red^*(\hat{x}) \cdot Red^*(\hat{y}) \cdot Red^*(\hat{z}) \cdot Wh(x, y, z, w)$
- (etc.)

Composite abstraction thus becomes *numerical abstraction*, the abstracting away from the *number* of something, only to retain the 'something'. The similarity between concepts differing only in this way is brought out clearly in the following expressions. The only variation is in the mereological subformula:

(27) $Red(x) \cdot Red(y) \cdot Same(x, y) \cdot Wh(x, y, w)$

(28) $Red(x) \cdot Red(y) \cdot Dis(x, y) \cdot Wh(x, y, w)$

Let us assume that we can take this idea further, beyond mere groups of quasi-atoms without further structure. If larger entities can be treated in this way, if more interesting types of interconnectedness can be represented, and once we start nesting numerical abstractions into each other, we shall be able to account for far more natural kinds of entities. For it seems to me that numerical abstraction is a pervasive aspect of the way human cognition chooses its categories. If X is a natural kind, say oak tree, then groups of X are a natural kind as well: two oak forests are similar in some respect, even though the number of trees may differ. It is not difficult to gather up examples of this: similarity of pattern (on clothing or wallpaper), the similarity between heaps of sand, or that between two samples of the same material.

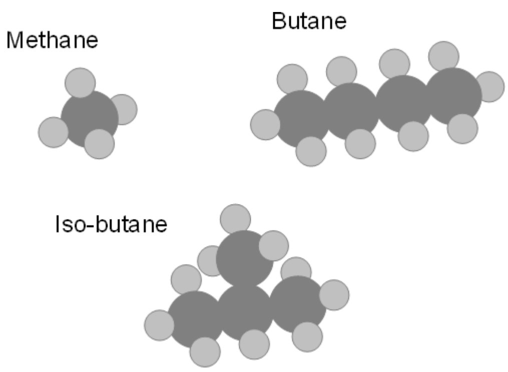
Chapter 6

Conclusion

In this thesis I have sketched the outlines of a theory providing a solution to the Identification Problem. The main part of the project has been to solve the problem for composite kinds, if it could be taken solved for elementary kinds. It has been solved, if the assumption is granted that *saturation* is a general characteristic of the way the world is structured. Then every naturalistic object, no matter how complex, can be categorized according to a variety of kinds, of sparse—natural—properties, if only its stock of basic properties and relations are complementary alternatives to one another. I believe that many aspects of the way we experience the world, ranging from everyday intuitions to the theories of elementary physics, give us reason to believe that this is so.

Instead of launching a theory about *structural universals*, my goal has been to describe what composite kinds are. For a nominalist there is no difference between the two, but most studies about structural universals are written at least with realist sympathies in mind; a position that has the virtue of sharpening the eye for the reality of sparse properties, but brings troubles as well. The case of *methane*, a chemical compound consisting of one carbon atom connected to four hydrogen atoms, is well-studied. Even though there is no particular issue about how these compounds relate in individual methane-atoms, for the universal things get utterly puzzling. There is only *one* carbon universal and *one* hydrogen universal. The universal methane should somehow comprise both, but how? Simply conjoining them will not produce a result that discriminates between methane (CH_4) and, say, butane (C_4H_{10}); and, as if to foil any hope that additional information about numbers of repeated instantiation would save the day, C_4H_{10} comes in two structural variants (isomeres): butane and iso-butane (Lewis, 1986a; Hawley, 2010). For the details

Figure 6.1: Organic molecules



see Lewis (1986a); the bottom-line is quite convincing: there is no easy way to make sense of methane as a structural universal.

There might of course be a *difficult* way to do so. Some authors have found it necessary to throw off the chains of traditional mereology to arrive at richer notions of parthood, able to deal with the problem just sketched. E.g. Mormann (2010) warns philosophers not to lag behind new insights with respect to mereology:

> For mathematicians the talk that "a group G has group H as a part three times over" is not empty at all but can be reconstructed as meaningful in terms of a general theory of structural parthood and composition that can be formulated in the general framework of category theory. (p 224)

Also Armstrong's *state of affairs* (e.g. Armstrong, 1986, 1997) need a 'non-mereological' sort of composition to allow them to be composed of both particulars and universals, without thereby generating lots of unintended overlaps between states of affairs.

My conclusion is different. I do not believe that there is anything wrong with classical mereology. To be sure, the kind *methane* can be straightforwardly rendered as a composite kind, only assuming saturation on the basic level. But chemical theory gives us exactly that! For the quasi-atomic kind *Atom* the alternatives are the elements, and every element has its own set of alternatives with respect to the primitive relation B (covalent bond). Hence, as we have seen, all the structurally different full concepts of this type are mergeable to monadic concepts, of molecules this time. Similarly the much-disputed difference between butane and iso-butane (both C_4H_{10}) is no problem. In the first compound the C-atoms lie in one row; in the second, three of them are bound to the remaining one. Clearly the configuration:

$$C(a) \cdot C(b) \cdot C(c) \cdot C(d) \cdot B(a,b) \cdot B(b,c) \cdot \underline{B(c,d)} \cdot (...H - bindings...)$$

(butane) differs from that of :

$$C(a) \cdot C(b) \cdot C(c) \cdot C(d) \cdot B(a,b) \cdot B(b,c) \cdot \underline{B(b,d)} \cdot (...H - bindings...)$$

(iso-butane). They are non-synonymous, hence alternatives for one another, and so are their monadic versions. Thus, *contra* Mormann (2012) representing this distinction does *not* require non-standard mereology.

To repeat, mereology is there to describe the extra-logical, *metaphysical* fact that common individuals show part-whole relations, and it appears to be successful in that—which is quite an achievement for a metaphysical theory. My observation is that it is *realism going beyond the most obvious individuals*, namely particulars like humans, kitchen knives and Higgs-bosons, that makes problems with this simple theory pop up instantly. If construed as entities, structural universals become a head-ache; as composite kinds, they are just kinds like any other.

I hope to have provided some reason to endorse the view that, if mereology assumes its proper place in metaphysics, it provides insights, even in its humble Boolean guise, that allow us to make big steps in solving the Identification Problem. And this, in turn, is invaluable for understanding what kind of structure reality has, and how we humans relate to it epistemically. Kinds, to this end, should not be reified, but they should be taken seriously!

Bibliography

[1] Armstrong, D.M. (1978) *A Theory of Universals: Universals and Scientific Realism, Vol. II.* Cambridge Univ. Press

[2] Armstrong, D.M. (1986) In defence of structural universals. *Australasian Journal of Philosophy*, 64 : 85-88

[3] Armstrong, D.M. (1991) Classes Are States of Affairs. *Mind*, 100 : 189-200

[4] Armstrong, D.M. (1997) *A World of States of Affairs*, Cambridge University Press, Cambridge

[5] Bar-Am, N. (2008) *Extensionalism: The Revolution in Logic*. Springer

[6] Bøhn, E.D. (2011) Commentary: David Lewis, Parts of Classes. *Humana.Mente Journal of Philosophical Studies*, 2011, 19 : 151–158

[7] Goodman, N. (1951) *The Structure of Appearance*, Bobbs-Merrill, New York

[8] Goodman, N. and W.V.O Quine (1947) Steps toward a Constructive Nominalism. *The Journal of Symbolic Logic* 12 : 105-122

[9] Hawley, K. (2010) Mereology, Modality and Magic. *Australasian Journal of Philosophy* 88 : 117–133.

[10] Herre, H. (2010) The Ontology of Mereological Systems: A Logical Approach. In: *Theories and Applications of Ontology*. Springer, 57-82

[11] Johnson, W.E. (1921) *Logic, Part I.* Cambridge University Press, Cambridge

[12] Jubien, M. (1993) *Ontology, Modality, and the Fallacy of Reference*. Cambridge University Press, Cambridge

[13] Kalhat, J. (2008). Structural universals and the principle of uniqueness of composition. *Grazer Philosophische Studien*, 76 : 57–77

[14] Kripke, S.A. (1981) *Naming and Necessity*. Blackwell

[15] Lewis, D. (1983) New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61 : 343-377

[16] Lewis, D. (1986a) Against Structural Universals, *Australasian Journal of Philosophy,* 62 : 25-46

[17] Lewis, D. (1986b) *On the Plurality of Worlds*, Oxford: Blackwell.

[18] Lewis, D. (1991) *Parts of Classes*, Oxford: Blackwell.

[19] Mormann, T. (2010) Structural Universals as Structural Parts: Toward a General Theory of Parthood and Composition. *Axiomathes* 20 : 209–227

[20] Mormann, T. (2012) On the mereological structure of complex states of affairs. *Synthese* 187 : 403–418

[21] Pagès, J. (2002) Structural Universals and Formal Relations. *Synthese*, 131-2 : 215-221

[22] Putnam, H. (1980) Models and Reality. *The Journal of Symbolic Logic* 45-3 : 464-482

[23] Quarfood, M. (2013) Interpretations of Kantian Disjunctive Judgment in Propositional Logic. In *Proceedings of the XI. International Kant Congress.* Berlin: De Gruyter

[24] Quine, W.V.O. (1969a) Ontological Relativity. in *Ontological Relativity and Other Essays.* Columbia Univ. Press

[25] Quine, W.V.O. (1969b) Natural Kinds. in *Ontological Relativity and Other Essays.* Columbia Univ. Press

[26] Rodriguez-Pereyra, G. (2002) *Resemblance Nominalism. A Solution to the Problem of Universals.* Clarendon Press, Oxford

[27] Sider, Th. (2007) Parthood. *Philosophical Review* 116 : 51–91

[28] Sinisi, V.F. (1983) Lesniewski's Foundations of Mathematics. *Topoi* 2 : 3-52

Appendix A

Mereological Predicates

Primitive predicates:

- $Wh(\mathbf{x}, w)$
- $Dis(x, y)$
- $Atom(x)$
- $All(x)$

Defined predicates:

- $Part(x, y) :\doteq Wh(x, y, y)$
- $Same(x, y) :\doteq Wh(x, y)$
- $Dis(\mathbf{x}) :\doteq \bigodot_{i,j \leq n, i \neq j} Dis(x_i, x_j)$
- $WD(\mathbf{x}, w) :\doteq Wh(\mathbf{x}, w) \cdot Dis(\mathbf{x})$
- $Exc(x, y) :\doteq Part(a, x) \cdot Dis(a, y)$ (excess)
- $PPart(x, y) :\doteq WD(x, a, y)$ (proper part)
- $POv(x, y) :\doteq Dis(a, b) \cdot WD(a, u, x) \cdot WD(b, u, y)$ (proper overlap)

Appendix B

Proofs of Chapter 4

Lemma 4.3. If something is distinct from something else, it is distinct from its parts:

$$Dis(x, y) \cdot Part(a, y) \Rightarrow Dis(a, x)$$

Proof.

$Dis(x, y) \cdot Part(a, y)$

$\quad \Rightarrow Dis(x, y) \cdot Part(a, y) \cdot Dis(a, x) + Dis(x, y) \cdot Part(a, y) \cdot Ov(a, x)$

Regarding the second disjunct:

$Dis(x, y) \cdot Part(a, y) \cdot Ov(a, x)$

$\doteq Dis(x, y) \cdot Part(a, y) \cdot Part(b, a) \cdot Part(b, x)$

$\quad \Rightarrow Dis(x, y) \cdot Part(b, y) \cdot Part(b, x)$

$(...) \Rightarrow Dis(x, y) \cdot Ov(x, y) \Rightarrow \bot$

So the alternative obtains, yielding the desired result.

□

Lemma 4.4. Parthood plus reverse excess makes proper parthood:

$$Part(x, y) \cdot Exc(y, x) = PPart(x, y)$$

where $PPart(x, y) :\doteq WD(x, a, y)$.

Proof.

$Part(x, y) \cdot Exc(y, x)$

$\doteq Wh(x, y, y) \cdot Exc(y, x)$

$\doteq Wh(y, x, y) \cdot Exc(y, x) \cdot Wh(y, x, w) \cdot WD(a, x, w)$ Axiom 9

$\doteq Same(y, w) \cdot WD(a, x, w)$

$\doteq WD(a, x, y) \doteq WD(x, a, y) \doteq PPt(x, y)$

□

Lemma 4.5. Parthood implies sameness or proper parthood:

$$Part(x, y) \Rightarrow Same(x, y) + PPart(x, y)$$

Proof.

$$Part(x,y) \doteq Part(x,y) \cdot y \cdot x$$
$$\Rightarrow Part(x,y) \cdot (Part(y,x) + Exc(y,x)) \qquad \text{Axiom 8}$$
$$(...) \Rightarrow Same(x,y) + PPart(x,y)$$

□

Lemma 4.6. Complements are unique:

$$WD(x,a,w) \cdot WD(x,b,w) \Rightarrow Same(a,b)$$

Proof.

$$WD(x,a,w) \cdot WD(x,b,w)$$
$$\Rightarrow WD(x,a,w) \cdot WD(x,b,w) \cdot (Part(a,b) + Exc(a,b))$$

Considering the second disjunct after expansion:

$$WD(x,a,w) \cdot WD(x,b,w) \cdot Exc(a,b)$$
$$\doteq WD(x,a,w) \cdot WD(x,b,w) \cdot Part(u,a) \cdot Dis(u,b)$$
$$\doteq WD(x,a,w) \cdot WD(x,b,w) \cdot Part(u,a) \cdot Part(a,w) \cdot Dis(u,b) \qquad \text{Lemma 4.1}$$
$$\doteq WD(x,a,w) \cdot WD(x,b,w) \cdot Part(u,w) \cdot Dis(u,b) \cdot Dis(u,x) \qquad \text{Lemma 4.3}$$

(To wit: $WD(x,a,w)$, so $Dis(x,a)$; plus $Part(u,a)$ makes $Dis(u,x)$.)

$$\Rightarrow Part(u,w) \cdot Dis(u,w) \Rightarrow \bot \qquad \text{Axiom 7}$$

Hence, the consequent must be $Part(a,b)$; but then likewise $Part(b,a)$, so $Same(a,b)$.

□

Lemma 4.7. If something is part of a whole but distinct from one constituent, it is part of the other:

$$Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x) \Rightarrow Part(a,y)$$

Proof.

$$Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x)$$
$$\Rightarrow Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x) \cdot (Part(a,y) + Exc(a,y))$$

The second disjunct:

$$Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x) \cdot Exc(a,y)$$
$$\doteq Wh(x,y,w) \cdot Part(a,w) \cdot Dis(a,x) \cdot Part(b,a) \cdot Dis(b,y)$$
$$\Rightarrow Wh(x,y,w) \cdot Part(a,w) \cdot Part(b,a) \cdot Dis(b,x) \cdot Dis(b,y) \qquad \text{Axiom 4.3}$$
$$(...) \Rightarrow Part(b,w) \cdot Dis(b,w) \Rightarrow \bot \qquad \text{Axiom 7}$$

□

Lemma 4.8. Mutual excess implies proper overlap or distinctness:

$$Exc(x,y) \cdot Exc(y,x) \Rightarrow POv(x,y) + Dis(x,y)$$

where $POv(x,y) :\doteq Dis(a,b) \cdot WD(a,u,x) \cdot WD(b,u,y)$.

Proof.

$Exc(x, y) \cdot Exc(y, x) \doteq Wh(x, y, w) \cdot WD(a, y, w) \cdot WD(b, x, w) \cdot Part(a, x) \cdot Part(b, y)$

Expanding the right-hand side by the expansion of the $Part$-segments will yield that the only consistent ones among the four disjuncts are:

$Wh(x, y, w) \cdot WD(a, y, w) \cdot WD(b, x, w) \cdot Same(a, x) \cdot Same(b, y)$ and

$Wh(x, y, w) \cdot WD(a, y, w) \cdot WD(b, x, w) \cdot PPart(a, x) \cdot PPart(b, y)$

The first directly implies $Dis(x, y)$. For the second only a few steps including the use of Lemma 4.6 are required to arrive at $POv(x, y)$.

□

Theorem 4.2.

$$x, y \Rightarrow Same(x, y) + PPart(x, y) + PPart(y, x) + POv(x, y) + Dis(x, y)$$

where $POv(x, y) := Dis(x, y) \cdot WD(a, u, x) \cdot WD(b, u, y)$ ("Proper Overlap").

Proof.

The desired result is obtained immediately by working out the expansion:

$x \cdot y \Rightarrow (Part(x, y) + Exc(x, y)) \cdot (Part(y, x) + Exc(y, x))$

□