

Chapter 11: Four Kinds of *Is_a* Relation

Ingvar Johansson

1. *General Introduction*

In many corners of the information sciences, including work on knowledge representation, description logics, and some object-oriented programming languages, the so-called *is_a* relation plays a prominent role. In this chapter, it will be argued that there are both material and formal reasons to distinguish between four kinds of *is_a* relation, which we will call (1) subsumption under a genus, (2) subsumption under a determinable, (3) specification, and (4) specialization.⁵¹

Genus-subsumption is the traditional way of creating classificatory trees of natural kinds; in particular, of creating the famous hierarchies of plants and animals in biology. However, it is also used in more practically oriented classifications of kinds such as citizens, patients, furniture, clothes, vehicles, and so forth. On the linguistic side, this kind of subsumption is usually mirrored, linguistically, by relations between nouns.

Determinable-subsumption, on the other hand, is not concerned with natural kinds but with qualities (properties) of different generality. For instance, as a determinate *scarlet* is subsumed under the determinable *red*. In such cases, we find relations between adjectives on the linguistic side.

Even though today, in the information and computer sciences, the two expressions ‘*a* is a *specification* of *b*’ and ‘*a* is a *specialization* of *b*’ are quite often used as synonyms of ‘*a is_a b*’, the terms ‘specification’ and ‘specialization’ will here be given more restricted meanings. In fact, these restricted meanings come close to the pre-computer world meanings of these terms. Whereas subsumptions are typically concerned with natural kinds and qualities, specifications and specializations are typically concerned with activities and processes. Prototypical specifications come out on the conceptual level primarily as relations between a verb-plus-adverb expression and a verb: ‘painting carefully’ is a specification of ‘painting’. Analogously, specializations come out primarily as relations between a verb plus a whole adverbial adjunct clause and a (usually

⁵¹ The *part_of* relation and the *instance_of* relation are not *is_a* relations at all, even though in natural languages one can say things such as ‘It *is a* part of the play’ and ‘He *is a* Swede’ (see Taivalsaari, 1996; Smith and Rosse, 2004; Smith, *et al.*, 2004).

transitive) verb which, when substantivized (converted into the language of nouns), gives: ‘painting a table’ is a specialization of ‘painting’. Of course, when the verbs are substantivized, nouns and adjectives can be used to represent specifications and specializations, too (see Figure 1).

In Figure 1, we have related some *is_a* expressions to some corresponding ordinary language sentences. In relation to this list, we will introduce a distinction between *the realist mode of speech* and *the conceptual mode of speech*, respectively. When a man in the street, or a scientist, asserts ‘a cat is a mammal’, he is talking about something that he believes to exist independently of his assertion. But when an information scientist writes ‘cat *is_a* mammal’, he may take himself to be talking only of concepts. The man in the street talks in the realist mode and the information scientist in the conceptual mode of speech; whereas the former may be said to look *through* concepts (and at the world), the latter may be said to look only *at* concepts (see Johansson, 2006). In everyday discourse, people switch from the realist to the conceptual mode when they are reading dictionaries and are reflectively translating between languages. The assertions ‘The German word ‘*Baum*’ means tree’ and ‘The German word ‘*alt*’ means old’ are assertions in the conceptual mode of speech. Each is in effect saying that a German and an English word have a concept in common. Assertions such as ‘*Dieser Baum ist alt*’ and ‘*This tree is old*’ belong to the realist mode of speech.

The distinction, now presented, has affinities with Rudolf Carnap’s classic distinction between the material and the formal mode of speech (in German: ‘*inhaltliche und formale Redeweisen*’) (1934). In fact, it can be looked upon as a version of Carnap’s distinction that has been freed from its original positivist-conventionalist setting and tied to a realist framework. The left column of Figure 1 can be read in both the conceptual and the realist mode of speech. The assertion ‘cat *is_a* mammal’ can be read either as ‘the concept of cat is a concept that is subsumed under the concept of mammal’ or as ‘the class of cats is subsumed under the class of mammals’. Note that, even though cats have (inherit) all the properties which mammals have in so far as they are mammals, the concept *cat* does not have all the properties that the concept *mammal* has. In *The Description Logic Handbook* (Baader, *et al.*, 2005, p. 5) it is asserted that:

The IS-A relationship defines a hierarchy over the concepts and provides the basis for the ‘inheritance of properties’: when a concept is more specific than some other concept, it inherits the properties of the more general one. For example, if a person has an age, then a woman has an age, too.

Figure 1: *Examples of is_a relations*

<i>is_a</i> expressions	Corresponding ordinary sentences about one individual case or about classes of cases	Corresponding ordinary sentences that are (at least seemingly) directly about universals
cat <i>is_a</i> mammal	a cat is a kind of mammal; cats are mammals	the cat is a mammal
mammal <i>is_a</i> animal	a mammal is a kind of animal; mammals are animals	
sailing ship <i>is_a</i> ship	a sailing ship is a kind of ship; sailing ships are ships	
ship <i>is_a</i> vehicle	a ship is a kind of vehicle; ships are vehicles	
scarlet <i>is_a</i> red	a scarlet thing is a kind of red thing; all scarlet things are red things	scarlet is a red hue
red <i>is_a</i> color	a red thing is a kind of colored thing; all red things are colored things	red is a color
running <i>is_a</i> activity	to run is to perform a kind of activity; all cases of running are cases of activity	running is an activity
painting <i>is_a</i> activity	to paint is to perform a kind of activity; all cases of painting are cases of activity	painting is an activity
careful painting <i>is_a</i> painting	painting carefully is a way of painting	careful painting is painting
house painting <i>is_a</i> painting	to paint a house is to paint (a certain kind of thing)	painting a house is painting
outside painting <i>is_a</i> painting	to paint an outside is to paint (a certain part of a thing)	painting an outside is painting
summer painting <i>is_a</i> painting	to paint in the summer is to paint (at a certain time of the year)	painting in the summer is painting

The quotation is understandable, but it conflates the realist and the conceptual mode of speech. Neither of the concepts ‘person’ or ‘woman’ has an age; but what can be referred to by means of these concepts, i.e., people and women, always have ages. As will be shown in what follows, in

order to become clear about the *is_a* relation in the conceptual mode of speech, one has to investigate some corresponding assertions that belong to the realist mode of speech. Sometimes, I will make it explicit when I switch between these modes of talking, but mostly I will trust that the context makes my mode of speech clear.

2. *Distinguishing between Sets and Classes*

The *is_a* relation can have as its relata real classes, concepts, or terms for concepts. Unfortunately, nowadays the terms ‘set’ and ‘class’ are often used as synonyms, but here they will be kept distinct (see Smith, *et al.*, 2005; Smith, 2005; Feibleman, 1974). Sets (as I will use this term) can be constructed by a simple act of will. Hence, a set can be created by means of artificial groupings (e.g., the set consisting of my neighbor’s cats together with his house), but no real class, such as the class of cats or the class of dogs, can be so delineated. A real class is a collection of entities that share a general language-independent feature (a universal or a type) in common. Such classes can be divided into two sorts (i) the extensions of universals (classes which consist of all and only the instances of some universal, for instance the class of all human beings), and (ii) defined (or partly fiat) classes, which are subclasses of extensions of universals delineated by means of some artificially created boundary (for instance: the class of all human beings in Leipzig).

Sets are identified by their members. The set of my neighbor’s three cats is the same as the set {Tibbles₁, Tibbles₂, Tibbles₃}, and this set remains the same even if he gives Tibbles₃ away to his daughter. Classes, in contrast, are identified by the universals and any fiat demarcations in terms of which they are defined. This means that, when considered with respect to time, classes (but not sets) can remain identical even while undergoing a certain turnover in their instances. Two distinct classes may have the same extension, but no distinct sets can have the same members. Hence, there is only one null set; but, in the sense of ‘class’ used here, the development of science forces recognition of several distinct zero-classes, i.e., classes that lack members. Famous examples of such classes from the history of science are phlogiston, planets that move around the Earth, and electron particles that orbit a nucleus.

To every non-zero class there is a corresponding set, but there is not a corresponding class for each non-zero set. In order to make this point more apparent, it will be helpful to compare subsumption schemas for classes

with those for sets. Let us take a look at a subsumption schema that consists of four levels of classes. It is only for the sake of expositional simplicity that each class in Figure 2 is divided into exactly two subclasses.

Figure 2: *A formal class subsumption schema*

Level 1	highest class: class A(1)							
Level 2	class A(2)				class B(2)			
Level 3	class A(3)		class B(3)		class C(3)		class D(3)	
Level 4	class A(4)	class B(4)	class C(4)	class D(4)	class E(4)	class F(4)	class G(4)	class H(4)

This schema for class subsumption must by no means be regarded as identical with the similar schema for set inclusion illustrated in Figure 3:

Figure 3: *A formal set inclusion schema*

Level 1	set A(1)							
Level 2	set A(2)				set B(2)			
Level 3	set A(3)		set B(3)		set C(3)		set D(3)	
Level 4	set A(4)	set B(4)	set C(4)	set D(4)	set E(4)	set F(4)	set G(4)	set H(4)

In relation to Figures 2 and 3 some of the things already said about the class-set distinction will become apparent. If none of the lowest classes of a subsumption schema is a zero class, then a corresponding set inclusion schema with the same number of sets as classes can always be constructed. One has only to regard the instances of each class as members of a corresponding set; however, the converse operation is not always possible. For example, let set A(4) be the set of cats that corresponds to the class of cats and let set B(4) be the set of red instances that corresponds to the class of red instances. The set A(3) is then simply the union of the sets A(4) and B(4), but there is no corresponding class A(3), because every class has to have some kind of internal coherence of the sort that is provided by a

common universal, but there is no such common universal between the class of cats and the class of red instances.

The following are some more examples that, like the cat and red example, will clarify the intuition behind such a conception. There is a *set* whose members consist of all temperature instances and of all mass instances, but there is no corresponding *class*. There is a set whose members consist of all molecules and of all cells, but there is no corresponding class.

In philosophy, much has been said about what, if anything, can constitute the kind of internal coherence (or unity) which distinguishes classes and sets (or, more generally, between natural kinds and arbitrary collections). The debate is still going on. The opposing positions are called realism, conceptualism, and nominalism. Realism entails that classes exist in reality independently of minds. Conceptualists hold that classes depend upon mental acts. Nominalists encourage us to believe that classes are simply that which is picked out by the use of general terms. The view put forward here might be called realist (there are completely mind-independent classes) with an admixture of conceptualism (some classes are partly fiat; more about this after Figure 5 below). However, for the purposes of this paper, it is enough if the reader accepts some conception of internal coherence that makes the distinction between class subsumption and set inclusion viable.

Let us summarize:

- from a semantic point of view, no class can be introduced or defined merely by means of an act of will, though many sets can;
- from an ontological point of view, sets can be identified with their members, but classes cannot; there can be only one zero set, but there can be many zero classes;
- from a temporal-ontological point of view, (i) certain kinds of sets can be tied to temporally located particulars; (ii) there are classes of activities and processes just as there are classes of objects and quality instances.

3. *Genus-Subsumption versus Determinable-Subsumption*

Classes of natural as well as artificial kinds (e.g., atoms, molecules, plants, animals, furniture, clothes, and vehicles) may stand in subsumption relations, but so may classes of qualities (e.g., colors, volumes, masses, and dispositional properties). As the class of cats is subsumed by the class of

mammals, which, in turn, is subsumed by the class of animals, so the class of scarlet instances is subsumed by the class of red instances, which, in turn, is subsumed by the class of color instances. With respect to individual things and spatiotemporal quality instances, these subsumptions imply:

- necessarily, if a certain particular is a cat then it is a mammal, and if it is a mammal it is an animal;
- necessarily, if there is an instance of being scarlet then there is an instance of being red, and if there is an instance of being red there is an instance of being colored;
- necessarily, if a certain particular is an animal, then it has to be an animal of a certain kind;
- necessarily, if there is an instance of being colored, then there is also an instance of some specific color hue.

Early in the twentieth century, the Cambridge philosopher W. E. Johnson (1921) argued against the view the two triple-subsumptions cat-mammal-animal and scarlet-red-color represent one single subsumption relation that relates different kinds of entities (natural kinds and qualities, respectively). Rather, he held, there are two different kinds of subsumption relation, forming genera-species hierarchies on the one hand and what Johnson termed determinable-determinate hierarchies on the other (see also Johansson, 2000). The reason for the assumption of one single subsumption relation is that both cases have, on the level of the corresponding extensions, the same class inclusion relation in common. The reason for the difference comes from the fact that species and genera (and all natural and artificial kinds of things) have monadic qualities by means of which they can be characterized, whereas determinate and determinable qualities cannot be so characterized. They are themselves qualities and so can only be characterized by means of their *similarity relations* to other qualities. The class of mammals can be defined as belonging to the *genus* animal, and as such having the specifically differentiating feature (*differentia specifica*) that the females are normally able to produce milk by means of which their offspring are first fed. The class of red instances cannot similarly be defined as colors (which would be the genus) that have in common a certain *differentia specifica* that is distinct from just being red. John Searle (1959, 143) describes this difference between species and determinates as follows:

A species is a conjunction of two logically independent properties—the genus and the differentia. But a determinate is not a conjunction of its determinable and some other property independent of the determinable. A determinate is, so to speak, an area marked off within a determinable without outside help.

If we define mammals (for short) as feeding-offspring-with-milk animals, then ‘feeding-offspring-with-milk’ and ‘animal’ are treated as being logically independent, i.e., they can neither be defined in terms of nor subsumed by each other. Even though there are no plants that produce milk, such plants are not logically impossible. One can then adequately say, with Searle, that mammals are marked off from other animals with outside help. But one cannot similarly mark off red from color (and scarlet from red) with outside help.

The need to distinguish between the genus-species distinction on the one hand and the determinable-determinate distinction on the other becomes even more obvious if we consider several subsumption levels simultaneously. Table 2 is a jointly exhaustive and pairwise disjoint (JEPD) subsumption schema that consists of four levels of classes which can be assumed to represent natural kinds of some sort. The classes on each level are mutually exclusive, and this entails that no class is subsumed by more than one class on the level above it. The schema ranges from a highest class (genus or determinable) via two intermediate levels to the lowest classes (species or determinates). All classes on the intermediate levels are species or determinates *in relation to* the higher and subsuming classes, but genera or determinables *in relation to* the lower and subsumed classes. Only the highest genus/determinable is a genus/determinable in a non-relative sense, and only the lowest species/determinates are species/determinates in this same non-relative sense.

For simplicity’s sake, we will abstract away from epistemology and ontological error, and talk as if all examples represent subsumption relations between non-empty classes. Genus-subsumption schemas represent the way pre-Darwinian biologists classified plants and animals, but such schemas are often used today outside of phylogenetic taxonomy. When a genus-subsumption taxonomy has become established, it can be used to lay down so-called Aristotelian real definitions, i.e., definitions that are primarily definitions of universals rather than of concepts. Philosophers who claim that *only* concepts can be defined are doing one of two things: they either (explicitly or implicitly) deny the existence of language-independent universals, or they restrict the term ‘definition’ in such a way that many definitions in the natural sciences cannot be called definitions.

Figure 4: *The formal structure of Aristotelian definitions of genera and species*

Class	Definition
class A(2)	$A(1) \cap a$
class B(2)	$A(1) \cap b$
class A(3)	$A(2) \cap c = A(1) \cap a \cap c$
class B(3)	$A(2) \cap d = A(1) \cap a \cap d$
class C(3)	$A(3) \cap e = A(1) \cap b \cap e$
class D(3)	$A(3) \cap f = A(1) \cap b \cap f$
class A(4)	$A(3) \cap g = A(1) \cap a \cap c \cap g$
class B(4)	$A(3) \cap h = A(1) \cap a \cap c \cap h$
class C(4)	$B(3) \cap i = A(1) \cap a \cap d \cap i$
class D(4)	$B(3) \cap j = A(1) \cap a \cap d \cap j$
class E(4)	$C(3) \cap k = A(1) \cap b \cap e \cap k$
class F(4)	$C(3) \cap l = A(1) \cap b \cap e \cap l$
class G(4)	$D(3) \cap m = A(1) \cap b \cap f \cap m$
class H(4)	$D(3) \cap n = A(1) \cap b \cap f \cap n$

In a complete system of Aristotelian definitions for any given domain, one would start from the highest genus and present, stepwise, the definitions of the lower classes until the lowest classes (species) have been defined. In each such step the subsuming class is divided into two or more subsumed classes by means of some differentiating quality requirements. The classic Aristotelian example is ‘man =_{def} rational animal’ – signifying that the subsumed class *man* is defined by means of a more general subsuming class (*animal*) plus a quality requirement, namely that the class man should have the quality *rationality* as its specific difference in relation to the other classes on the same level. The definitional route just described is also used in much computer programming. (For a formally more detailed exposition of Aristotelian definitions and of relations between species and genera, see Smith, 2005. In Figure 4, some features of importance for this paper are highlighted.)

For the sake of simple exposition, we will introduce symbols for what might be called class intersection (\cap) and class union (\cup), respectively. Though the notions of intersection and union belong to set-theory and are purely extensional, this usage is not intended to indicate that we define a class in terms of its extension. It merely means that instead of ‘man =_{def} rational animal’, we can write ‘man =_{def} rational \cap animal’; and instead of ‘red =_{def} dark red or light red’ we can write ‘red =_{def} dark red \cup light red’. Let us now assume that we have one highest genus, and fourteen quality

classes (a, b, c, ..., n), one for each *differentia specifica*. All the classes, except the highest one, can then be defined, as in Figure 4.

From a purely definitional point of view, all the classes from A(2) to H(4) become classes of natural kinds, not classes of qualities, only because the highest class A(1) is a natural kind. By definition, if the presumed specific differences do not give rise to mutually exclusive classes, they cannot be called *differentia specifica*. In definitions like these the highest genus, as well as all the species-differentiating qualities, has to be – in relation to the subsumption schema – undefined (Berg, 1983). As is easily seen in Figure 4, the lower classes have (inherit) all the qualities that are essential to the classes above them.

Aristotelian definitions are advanced in the realist mode of speech. If we switch to the conceptual mode of speech, the given definitions of the classes turn into definitions of the corresponding concepts. If, in the course of scientific development, a specific taxonomy is revised, then new real definitions have to be substituted for the old ones. When this happens, it is often the case that new or partly new concepts have to enter the scene. When, for instance, it was discovered that the class of whales should not be subsumed under the class of fishes but under the class of mammals, the concepts of both whale and fish had to be redefined (Johansson, 1986).

When Figure 2 is used to represent subsumptions under determinables such as length, color, and mass, the following should be noticed. If one wants to use the schema as a basis for definitions, one cannot proceed as in the case of genus-subsumptions. This is because: (i) trivially, one cannot create divisions of a class only by means of the class itself, and (ii) since the highest class is now a determinable, there are no qualities external to the class that can create subsumed classes. Therefore, the only way possible is to define the higher classes by means of the lower ones; which means that the lowest ones have to be regarded as undefined in relation to the schema. Since the lowest classes do not overlap, the definitions of the higher classes have to be made by means of the operation of union (\cup). The schema in Figure 2 can then be used to make the definitions stated in Figure 5.

Figure 5: *The formal structure of definitions of determinables by means of determinates*

Class	Definition
class A(3)	$A(4) \cup B(4)$
class B(3)	$C(4) \cup D(4)$
class C(3)	$E(4) \cup F(4)$
class D(3)	$G(4) \cup H(4)$
class A(2)	$A(3) \cup B(3) = A(4) \cup B(4) \cup C(4) \cup D(4)$
class B(2)	$C(3) \cup D(3) = E(4) \cup F(4) \cup G(4) \cup H(4)$
class A(1)	$A(2) \cup B(2) = A(4) \cup B(4) \cup C(4) \cup D(4) \cup E(4) \cup F(4) \cup G(4) \cup H(4)$

Some observations on the set-class distinction may be of relevance here. If the definitions given would be definitions of sets instead of classes, then it would be tautologically and vacuously true that the union of A(4) to H(4) exhausts the set A(1); but if we are dealing with classes, however, then the highest determinable has to ensure that there is an internal coherence among the lowest determinates. Otherwise the latter would not be able to be subsumed under the *class* A(1). Therefore, the definition of the *class* A(1) as the union of the classes A(4) to H(4) is in effect a statement (non-vacuously true or false) that says that the members of the classes A(4) to H(4) jointly exhaust the class A(1). In those cases where the highest determinable and the lowest determinates, but no classes in between, are naturally pre-given classes (which we think is a very important sort of case (Johansson, 2000)), then all the in-between classes can be created by means of conventions. We then get a number of partly fiat classes, for which the conventionality in question is bounded by one bona fide class at the top of the subsumption schema and many bona fide classes at the bottom.

When fiat classes of the kind mentioned are created on some given level, then one can, in principle, let these classes be either overlapping or mutually exclusive, but systems with mutually exclusive classes function much better from a communicative point of view; they simply contain more information. Then, for instance, one knows for sure that if one person says ‘this is an A(3)’ and another person says ‘this is a B(3)’, both cannot be right.

In everyday life, we divide length instances into classes such as *very short*, *short*, *medium*, *long*, and *very long*; temperature instances are similarly divided into classes such as *very cold*, *cold*, *neither cold nor warm*, *warm*, and *hot*. Classes like these can both subsume more determinate classes as well as be subsumed under even broader classes. In

physics, the same determinables ground linear scales. Such scales are special cases of determinable-subsumption. In themselves, they contain only two levels, the level of the highest determinable (length and temperature, respectively) and the level of the lowest determinates. The latter level contains (is the union of) infinitely many classes, one corresponding to each real number. For instance, the concept of 5.000789000 m refers to one class of length instances, and the concept of 74.67823000 m refers to another class. In all probability, many such classes are zero classes.

One difference between genus-subsumption and determinable-subsumption can now be summarized as follows: definitions based on determinable-subsumptions have to move bottom up with the help of the operation of class union, whereas definitions based on genus-subsumptions can also move top down with the help of the operation of class intersection.

Both kinds of *is_a* subsumption relation distinguished so far must be kept distinct from another relation that is also sometimes called ‘subsumption’, namely the relation between an individual (particular) and a class. To keep them distinct, the latter relation should be called ‘instantiation’ or ‘instance_of’. Hopefully, an example is enough to make the distinction clear. If Pluto is a brown dog, then both the statements ‘Pluto instance_of dog’ and ‘Pluto instance_of brown’ are true, but the statements ‘Pluto *is_a* dog’ and ‘Pluto *is_a* brown’ are misnomers.

4. *Specification*

Is_a relations such as ‘careful painting *is_a* painting’, ‘careless painting *is_a* painting’, ‘fast painting *is_a* painting’, and ‘slow painting *is_a* painting’ seem to conform neither to what is typical of genus-subsumption nor to what is typical of determinable-subsumption. We will call them *specifications*. Let us explain. The class *careful painting* is not identical with the intersection of two logically independent classes: *painting* and *careful*. There is no class carefulness that exists as an independent entity. Carefulness is always *careful activity of some sort*. Furthermore, the carefulness in ‘careful painting’ is distinct from the carefulness in ‘careful reading’, ‘careful cleaning’, ‘careful watching’; each of these carefulnesses is logically secondary to, and takes part of its essence from, the kind of activity that is in each case mentioned. Therefore, careful painting cannot be subsumed under painting as a species is subsumed under a genus. And what goes for *careful painting* goes for *careful painter*, too. It is a well

known fact in philosophy, linguistics, and the information sciences that (to talk in the conceptual mode of speech) the extensions for expressions such as ‘being a careful painter’, ‘being a fast painter’, and ‘being a good painter’ cannot be analyzed as the intersections of the extension of the expression ‘being a painter’ with the extensions of the expressions: ‘being careful’, ‘being fast’, and ‘being good’, respectively.

The difference between specification and determinable-subsumption is not equally clear, but one aspect of this distinction is the following. In the way we have shown, determinable-subsumption allows for definitions by means of the union of the subsumed classes, but it seems impossible to define any activity as the union of a number of specifications. For instance the extension of ‘painting’ cannot be regarded as identical with the union of the extensions of ‘careful painting’, ‘careless painting’, ‘fast painting’, ‘slow painting’, and so on for all possible specifications. Unlike genus-subsumptions and determinable-subsumptions, specifications cannot ground definitions at all.

The general remarks made above in relation to activities can be repeated in relation to processes (e.g., burning, digesting, and circulating). However, it has to be noted that some possible specifications of activities (e.g., careful and careless) cannot be specifications of processes, whereas others (e.g., fast and slow) are possible as specifications of both activities and processes. Specifications differ in structure from both genus-subsumptions and determinable-subsumptions, but it is easy to conflate them, and it is especially easy to conflate specification with determinable-subsumption. The distinction is nonetheless reflected in everyday language. We say that ‘painting is a *kind* of activity’ but that ‘painting carefully is a *way* of painting’. The crux of the matter is that different activities are not specifications but determinates of ‘activity’. That is, ‘painting’ is a determinate that is determinable-subsumed by ‘activity’, whereas ‘careful painting’ is a specification of ‘painting’; similarly, ‘careful activity’ is a specification of ‘activity’. This complication can create a need to combine in one and the same classificatory tree both determinable-subsumptions (painting → activity) and specifications (careful painting → painting).

The relation of specification seems not to be confined to activities and processes. Whereas (consciously perceived) color hues obviously are determinable-subsumed under the class of (consciously perceived) colors, the same is not true for color intensities and degrees of color-saturation. They seem to be specifications of color hues just as carefulness is a specification of activities. When two different color hues, say a determinate

red and a determinate blue, have the same intensity (or degree of saturation), the intensity (saturation) is logically secondary to, and takes part of its essence from, the color hue in question; not the other way round. The fact that color hues are determinates but color intensities and saturations are specifications is quite compatible with the fact that color hue, color intensity, and color saturation can, as in the Munsell color solid, be ordered along three different dimensions in an ordinary picture or in a three-dimensional abstract space (compare Figure 8, which combines a subsumption relation with one specification relation).

5. *Specialization*

Here are some examples of *is_a* relations that are *specializations*: ‘house painting *is_a* painting’, ‘outside painting *is_a* painting’, ‘summer painting *is_a* painting’, ‘car driving *is_a* driving’, ‘food digesting *is_a* digesting’, and ‘paper printing *is_a* printing’. In these cases, the class on the left of the *is_a* relation does not *specify* the activity mentioned on the right; it is doing something else. It relates the activity mentioned on the right to something (houses, outsides, and summers) that exists completely independently of this activity. This fact makes it immediately clear that specializations cannot possibly ground definitions of the activities that they are specializing. As we normally use the concept of specialization, we can say that one painter has specialized in painting houses and another in painting chairs, one in painting outsides of houses and another in painting insides. This is our main reason for having chosen the label ‘specialization’. However, our choice is in conformity with the terminology of a paper that has previously mentioned the feature that I am now trying to make even more clear; the author in question talks about *specializing criteria* as a certain kind of subsumption (*is_a*) principle (Bernauer, 1994).

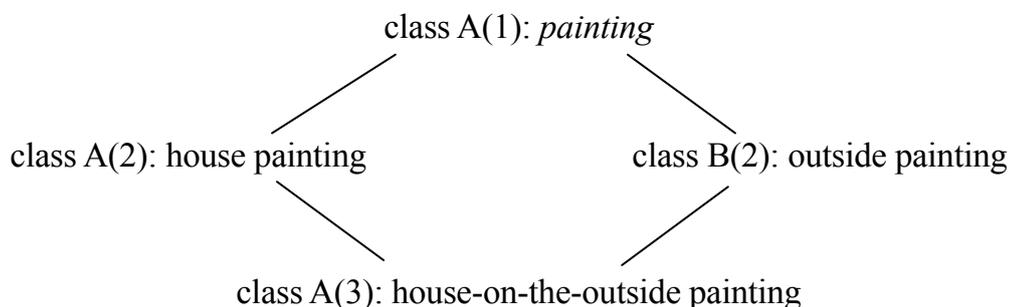
Some activities are simply activities performed by a subject (e.g., swimming and running), whereas others also involve one or several objects that are acted upon (e.g., painting a house and driving a car). Similarly, some processes simply occur in an object (e.g., rusting and burning), whereas others involve also one or several objects that the process in question acts upon (e.g., digesting food and printing papers). It is only in the *acting-on* kind of cases that specialization of activities and processes can come about. When there is talk about painting, driving, digesting, and printing as such, one knows that there is an object that has been abstracted away. It is this missing object that re-enters when a specialization is

described, or when a corresponding *is_a* relation is stated. Nothing like this occurs in subsumptions and specifications.

In all of the examples used above, the specializations are described by means of transitive verbs (or substantivizations of such verbs). And this is no accident. Transitive verbs are defined as verbs that can take (and often require) an object, whereas intransitive verbs cannot. Nonetheless, even intransitive verbs admit of specializations. This happens when the activity (process) in question becomes related to a certain kind of time period or a certain kind of place: ‘summer swimming *is_a* (specialization of) swimming’ and ‘pool swimming *is_a* (specialization of) swimming’.

Normally, an activity can be specialized in several different directions. One can paint a house, a car, or whatever. As soon as the object painted is such as to have both an outside and an inside, one can paint either the one or the other. Similarly, one may paint at a certain time of the year or at a certain kind of place. Therefore, some specializations have to have more than one *is_a* relation to the next level. Let us specialize ‘painting’ along two different directions: what kind of object that is painted and which part of an object that is painted. This, then, yields the following *is_a* schema:

Figure 6: *A double-specialization schema*



In words: house-on-the-outside painting *is_a* house painting; house-on-the-outside painting *is_a* outside painting; house painting *is_a* painting; and outside painting *is_a* painting.

6. *Single and Multiple Inheritance*

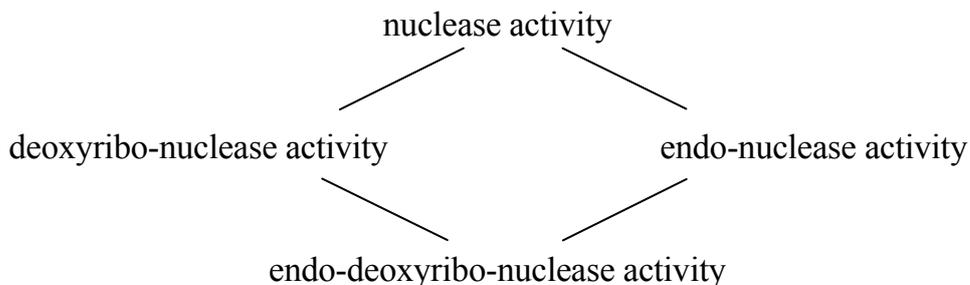
In the distinction between single and multiple inheritance, the concept of inheritance seems originally to have referred to inheritance of qualities in genus-subsumptions. A subsumed genus inherits all the properties that are essential to the subsuming classes. If a certain genus is subsumed by only one class on the nearest upper level, then there is single inheritance of qualities. If it is subsumed by two or more genera, then there is multiple inheritance. In determinable-subsumptions there are no real inheritances of

qualities apart from the inheriting of the highest determinable; the rest is, as we have explained, a matter of mere unions of the lowest determinates. Specifications, too, do not involve any literal quality inheritances. Nonetheless, the distinction between single and multiple inheritance is sometimes applied to each kind of *is_a* relation that we have distinguished. This means that when, in this general sense, it is stated that there is multiple inheritance, it is merely stated that the left-hand class of an *is_a* relation has some *is_a* relation to more than one class on the next level up.

From our remarks on genus-subsumption and determinable-subsumption, it follows that in both cases the default norm for such *is_a* hierarchies should be that they contain no multiple inheritances. With respect to specification, however, it does not even make sense to speak about multiple inheritance *of only specifications*. As we have analyzed ‘careful painting’, it can only have a specification relation to ‘painting’, since ‘careful’ in ‘careful painting’ has no complete meaning independently of painting. With respect to specializations, however, things are completely different. Here we get multiple inheritances as soon as there are two or more different directions that a specialization can take. In Figure 5, ‘house-on-the-outside painting’ is multiply (doubly) inherited.

Multiple inheritances are consciously and, according to my analysis, correctly used in the Gene Ontology (see *Gene Ontology*). The Gene Ontology Consortium states that ‘GO terms are organized in structures called directed acyclic graphs (DAGs), which differ from hierarchies in that a “child” (more specialized term) can have many “parents” (less specialized terms)’ (Gene Ontology Consortium, ND). The GO consortium uses the concept of ‘specialization’ as a synonym for ‘*is_a* relation’, but whenever a child in a GO graph has more than one *is_a* parent, then at least one of the *is_a* relations in question is a specialization in my restricted sense.

Figure 7: *A specialization schema with examples from the Gene Ontology*

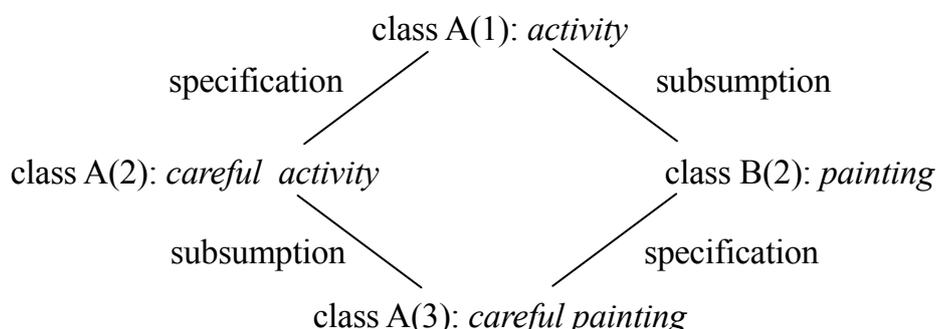


Let us exemplify. In the GO ontology for molecular functions one finds (hyphens added) ‘endodeoxyribo-nuclease activity’ (GO:0004520) inherited from both ‘deoxyribo-nuclease activity’ (GO:0004536) and ‘endo-nuclease activity’ (GO:0004519); both of the latter are, in turn, inherited from ‘nuclease activity’ (GO:0004518). Setting the arrows of GO’s graphs aside, these specializations can be represented as in Figure 7.

This is one of numerous examples of specializations that can be extracted from the Gene Ontology. A nuclease activity is an activity (performed by an enzyme) that catalyzes hydrolysis of ester linkages within nucleic acids. Such activity can be specialized along at least two different directions: (i) according to *what* is acted on (deoxyribonucleic acid, DNA, or ribonucleic acid, RNA), and (ii) according to *where* the action takes place, i.e., cleaving a molecule from positions inside the molecule acted on (‘endo-’), and cleaving from the free ends of the molecule acted on (‘exo-’), respectively. Since nothing stops the specialization from going in both directions at once, we get the schema for ‘nuclease activity’ in Figure 7, which is completely analogous to the schema for ‘painting’ in Figure 6.

Specializations allow and often require multiple inheritance. They differ in structure from genus-subsumptions, determinable-subsumptions, and specifications. However, we have so far spoken of hierarchies or graphs consisting of only one of these kinds of *is_a* relation, but the different *is_a* relations can also be combined with each other. (They can also, as in the GO, be combined with the *part_of* relation.) And in such mixed cases, too, multiple inheritance can be the normal and the required kind of inheritance. Two examples may show what we mean.

Figure 8: A combined specification and determinable-subsumption schema

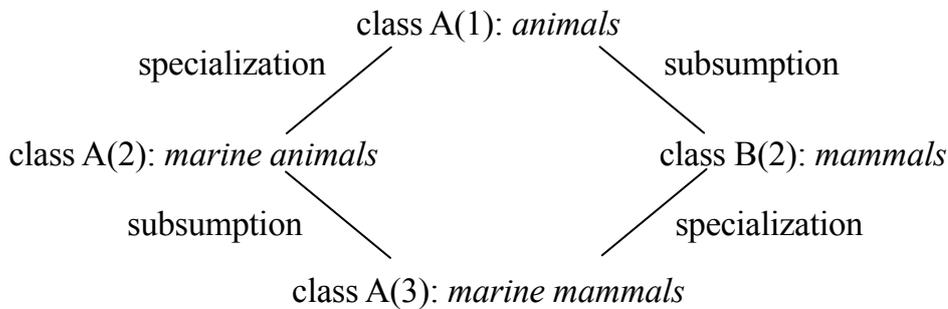


In the first example, Figure 8, ‘careful painting’ is doubly inherited. On the left hand side, the *is_a* relation is one of subsumption, but on the right

hand side it is one of specification. When the whole Figure is taken into account, symmetry is displayed. Two specifications are diagonally opposed, and so are two subsumption relations.

In traditional non-phylogenetic classifications of animals the *differentia specifica* are (broadly speaking) properties inhering in the organisms. But, of course, one can also try to classify animals according to where, when, and on what they perform various activities. Some live on land and some in the sea; some sleep during the night and some in the day; some eat meat and some do not. Therefore, classes of animals can via their activities also be made relata in specialization relations. Most mammals live on land but whales live in the sea. We may speak of a class *marine mammals* that can be placed in an *is_a* schema such as that of Figure 9.

Figure 9: A combined specialization and genus-subsumption schema



We have sometimes mentioned the ‘instance_of’ relation (e.g., at the end of section 2). Now, in order to avoid all misunderstandings, we need to do it again. Everything that has been said about multiple inheritance above relates to *is_a* relations and not to ‘instance_of’ relations. Trivially, an individual can instantiate many classes and in this special sense have multiple inheritance (better: ‘multiple instantiation’) when placed in a slot in a matrix. Many matrices that are used in the social sciences and in epidemiology to display correlations have this character. A simple but fictive example that contains this kind of multiple (double) inheritance for a group of hundred persons is presented in Figure 10 (from Asplund, 1968).

Figure 10: A correlation matrix relating political views to political interest

	Republicans	Democrats	Independents
High	14(persons)	16	5
Medium	19	17	7
Low	5	5	12

Here, each of the fourteen individual persons in the upper left slot inherits two features: having high political interest and being Republicans. Such tables must by no means be conflated with those in Figures 6 to 9 above. Note, though, that if an individual is an instance of a certain class, then he is automatically also an instance of all classes that subsume this class.

7. *Philosophy and Informatics*

Can the taxonomy of *is_a* relations presented in the above be of any use in informatics? Let us answer by means of a detour.

No observations can be reported, and no reasoning can take place, without classification. But a classification is not necessarily a taxonomy: it need not be a *systematized* classification. During medieval times, alchemists made extensive classifications of substances, and herbalists did the same with respect to plants, but in neither case was a real taxonomy created. In a sense, the alchemists and the herbalists were too practically minded. But, things changed with the advent of modern chemistry and botany. Remarkable taxonomies appeared with remarkable repercussions on scientific development.

Today, information scientists help other scientific disciplines, as well as practical endeavors of all kinds, to systematize their respective classifications. But, curiously enough, they nonetheless seem to have no deep impulse to systematize their own use of various kinds of *is_a* relations and different kinds of definitions. Despite being a philosopher by trade, I dare to assert that at least some information scientific work can be done more efficiently if those involved would (a) accept that there is a distinction to be drawn between sets and classes, and (b) become aware of the tetrachotomy of *is_a* relations that we have put forward. Thus far, much that has been created, on a purely pragmatic basis, has rested on principles that were made explicit only later. However, once discovered, such principles can be consciously put to use and, thereby, make future similar work simpler and more effective. Without any explicit talk of that special kind of *is_a* relation which we have called specialization, the authors of the Gene Ontology chose to work with directed acyclic graphs instead of the set-theoretical inclusion relation, but this fact is no reason not to make the next generation of information scientists aware of the existence of the different kinds of *is_a* relations thereby involved.

