Cite as: Bergman, M. K. A KR Terminology. in *A Knowledge Representation Practionary: Guidelines Based on Charles Sanders Peirce* (ed. Bergman, M. K.) 129–149 (Springer International Publishing, 2018). doi:10.1007/978-3-319-98092-8_7

Official site: https://link.springer.com/book/10.1007/978-3-319-98092-8

Full-text: http://www.mkbergman.com/publications/akrp/chapter-7.pdf

Abstract: Peirce's concept of prescission captures the most fundamental expression of a hierarchical relationship, stated as the relation, prescind. When paired with the lessons of prior chapters, we end up with an expressive grammar for capturing all kinds of internal and external relations to other things.

7

A KR TERMINOLOGY

Speculative grammar, the theory of the nature and meaning of signs, is the first of the three branches of logic according to Peirce. The basic idea of a speculative grammar is simple when applied to a new concept or domain, such as knowledge representation. What is the terminology — vocabulary and relations — that may be involved in understanding the questions or concepts at hand? What is the 'grammar' for relating this terminology to the logics we need to help increase our understanding of the domain? How shall we split and organize these concepts? That is, how shall we categorize our domain? How do we combine these elements into assertions and statements and then test them for truth and accuracy? Through this grammar, in the KR context, can we maximize the structural features within our focus of inquiry useful to machine learners?

The term 'semiosis' most often brings to mind Peirce's theory of signs.[1] However, for Peirce semiosis was a broader construct still, representing his overall theory of logic and truth-testing. Signs, symbols, and representation are the first part of this theory, the speculative grammar about how to formulate and analyze logic. Though he provides a unique take on it, Peirce's idea of speculative grammar, which he ascribed to Duns Scotus, perhaps should be traced back to the 1300s and the writings of Thomas of Erfurt, one of the so-called Modists of the medieval philosophers.[2] Here is how Peirce placed speculative grammar within his theory of logic:

> "All thought being performed by means of signs, logic may be regarded as the science of the general laws of signs. It has three branches: (1) *Speculative Grammar*, or the general theory of the nature and meanings of signs, whether they be icons, indices, or symbols; (2) *Critic*, which classifies arguments and determines the validity and degree of force of each kind; (3) *Methodeutic*, which studies the methods that ought to be pursued in the investigation, in the exposition, and in the application of truth." (1903, CP 1.191, EP 2:260)

In terms of the logic triad, speculative grammar is thus a Firstness in Peirce's category structure. Firstness is meant to capture the possibilities of the domain at hand. Secondness is meant to capture the particular facts or real things of the domain at hand, the *critic* in terms of the logic triad. Thirdness is meant to capture methods for discovering the generalities, laws or new knowledge arising from the domain, the *methodeutic* branch of the triad. We may apply this construct to any topic, from signs to logic and science. The 'surprising fact,' or new insight arising from Firstness or Thirdness, points to potentially new topics that may themselves become new targets for this logic of semiosis.

Without the right concepts, terminology, or bounding — that is, the speculative grammar — it is impossible to understand or compose the *objects* (conceptual and material) within Secondness that populate the domain at hand. Without the right language and concepts to capture the connections and implications of the domain at hand — again, part of its speculative grammar — it is not possible to discover the generalities or the 'surprising fact' or Thirdness of the domain. The speculative grammar is thus needed to provide the right constructs for describing, analyzing, and reasoning over the given domain. Our logic and ability to understand the focus of our inquiry require that we describe and characterize the domain of discourse with proper scope and relationships. How well we bound, characterize, and signify our problem domains — again, the speculative grammar — directly relates to how well we can reason and inquire about that space.

Let's take a couple of examples to illustrate this. First, imagine van Leeuwenhoek first discovering 'animalcules' under his early microscope. Over the ensuing years, new terms and concepts like flagella, cells, and vacuoles were coined and systematized to enable a further understanding of microorganisms, requiring careful inspections and consensual vocabulary. Second, imagine 'action at a distance' phenomena such as magnetic repulsion or static electricity causing hair to stand on end. For centuries these phenomena were assumed caused by atomistic particles too small to see or discover. Only when Hertz was able to prove Maxwell's equations of electromagnetism in the mid-1800s were the concepts and vocabulary of waves and fields sufficiently developed to begin to unravel electromagnetic theory in earnest. Progress required the right concepts and terminology.

For Peirce, the triadic nature of the sign — and its relation between the representamen, its object, and its interpretant — was the speculative grammar breakthrough that then allowed him to better describe the process of sign-making and its role in the logic of inquiry and truth-testing (semiosis). Because he recognized it in his work, Peirce understood a conceptual grammar appropriate to the inquiry at hand is essential to further discovery and validation. As Peirce says in his first paper outlining his early logic of relatives:[3]

> "The fundamental principles of formal logic are not properly axioms, but definitions and divisions; and the only facts which it contains relate to the identity of the conceptions resulting from those processes with certain familiar ones." (1870, CP 3.149)

We begin our analysis of a speculative grammar suitable to knowledge representation with the relevant 'things' (nouns) that populate our world and how we organize them in relation to one another.[4] We then expand our discussion of *relations* to include *actions* and perceptions (verbs) between these things, as well as how we talk about or describe those things. Things and relations combined enable us to make *statements* and *assertions*. In the aggregate, multiple statements interact to create many kinds of information structures, some with impressive analytical properties discussed in later chapters. In knowledge representation, terminology can be a tricky business, since different approaches to KR adopt different terms, sometimes overlapping or in conflict with other schemes. I try to point out some of these conflicts for key terms in the three chapters of this *Part II* on grammar. Throughout this chapter — indeed, this book — we *italicize* the basic KR terminology we have adopted for this book and KBpedia. The *Glossary* consolidates this terminology in one location.

THINGS OF THE WORLD

We watch our children first learn the names of things as they begin mastering language. The learning focus is on nouns and building a vocabulary about the things that populate the tangible world. By the time we start putting together our first sentences, typified in books such as Dick and Jane and the dog Spot, our nouns get increasingly numerous and rich, though our verbs remain simple. We acquire terms in our early language more about different kinds of objects than different kinds of actions (though concepts such as 'More' or 'Want' or gesturing to the mouth to signify 'Eat' are learned early!). Our early verbs are fewer in number and much less varied than the differences of form and circumstance we can see from objects. Our knowledge artifacts reflect this imbalance.

Entities, Attributes, and Concepts

Entities and concepts dominate most knowledge graphs. For example, knowledge base constituents of KBpedia, such as Wikidata, Wikipedia or GeoNames, have millions of concepts or entities within them, but fewer than a few thousand predicates (approx. 2500 useful in Wikidata and 750 or so in DBpedia and schema.org).

Entities are the individual, real things in our domain of interest; they are nameable things or ideas that have an identity, are defined in some manner, can be referenced, and may be related to *types*. Entities are most often the bulk of an overall knowledge base. An entity is an individual object or *instance,* a Secondness, of a *class*, a Thirdness. When affixed with a proper name or label, we term it a *named entity* (thus, named entities are a subset of all entities). *Attributes* describe and characterize entities. We connect or relate entities to one another through *external relations*. How we refer to, signify or index these things is what we call *representations*. An entity may be independent or separate or can be part of something else, such as parts of a whole. Entities cannot be topics or types or datatypes.[5]

We look to separate the existence of some things different from other things by the nature of their characteristics, what we can observe and describe for that given thing. So, we describe shapes, sizes, weights, ages, colors and characteristics of things with increasingly nuanced vocabularies. We note that grasses have linear or simple leaves, oaks have serrated or wavy-shaped leaves, and carrots have branched or compound leaves. We distinguish hair color, eye color, place of birth, current location and a myriad of factors. Each one of these factors becomes an *attribute* for that object, with the specific values (simple *v* wavy *v* compound) distinguishing instances from one another. We can also assign *values* to attributes, such as having an age of 7 years or a height of 120 cm. Attributes do not exist independently from the things they characterize.[*] For example, 'round' or 'blue' are not things unto themselves but are modifiers or qualifiers or characteristics of particular things. Attributes in Peirce's universal categories are a Firstness. Chen described similar entity-attribute distinctions in his attempt to find common ground across the network, relational and entity set models in today's commonly used E-R model.[6]

A *concept is* something we conceive in the mind, such as an idea or a grouping of like things. When we organize these things according to their shared and natural attributes, a topic we discuss in more detail in *Chapter 10*, we call them a *type*.[7] Concepts and types are not discrete, tangible things, but are constructs of thought. Topics are a form of a concept, but as used herein are more of a complex of concepts, ideas, and entities. Note this use of topic contrasts to that used in *topic maps*, which subsumes the terms of concepts, entities, and events used herein.[8] Concepts and types are *generals*, a Thirdness in Peirce's universal categories.

Here are some other terms you may encounter in other grammars or knowledge representations for these terms:

KBpedia Terminology	*Terminology Used Elsewhere*	
entity	▪ object ▪ instance ▪ exemplar ▪ element ▪ particular	• member • record • individual • dependent variable • token
attribute	• property • predicate • relationship • feature • facet	• dimension • characteristic • field • header • independent variables
type	• concept • kind • set	• collection • type • class

Table 7-1: Comparison of Common Noun Terms

The distinctions between entities and concepts are often murky in the real world. For example, let's consider the 'toucan' bird, which we may refer to by word or pic-

* Though Peirce, as do we, came to believe that Firstness (and Thirdness, for that matter) was real. For something to exist, it must be actual, which is Secondness.

ture. When we inspect what might be a description of a toucan on Wikipedia, we see that the term more broadly represents the family of *Ramphastidae*, which contains five genera and forty different species. The picture we use to refer to toucan may be, say, that of the keel-billed toucan (*Ramphastos sulfuratus*). However, if we view the images of a list of toucan species, we see just how physically divergent various toucans are from one another. Across all species, average sizes vary by more than a factor of three with great variation in bill sizes, coloration, and range. Further, if I assert that the picture of the toucan is that of my pet keel-billed toucan, *Pretty Bird*, then we can also understand that this representation is for a specific individual bird, and not the physical keel-billed toucan species as a whole. The point is not a lesson on toucans, but an affirmation that distinctions between what we think we may be describing occur over multiple levels. The meaning of what we call a 'toucan' bird is not embodied in its label or even its name, but in the accompanying referential information that places the referent into context. Without such accompanying context, the standalone word or picture of 'toucan' may represent either an individual entity or one of perhaps multiple types. I discuss further the importance of context and 'things, not strings' in *Chapter 10*.

What is an Event?

Events are like entities, except they have a discrete time beginning and end. Are events entities, and, if not, what are they? Events are part of time, occupy some length of time, and sometimes are so notable as to get names, either as types or named events, such as *germination* or *World War II*. Events are undoubtedly different from tangible objects which occupy some space, have physicality, exist over some length of time, and also get names as types or named instances. Moreover, both of these are different still than concepts or ideas that are creatures of thought. How to place the notion of events within a consistent worldview is one test for the coherency of a given knowledge representation.

The philosophical question of *What is an event?* is readily traced back to Plato and Aristotle. One place to start is the *Stanford Encyclopedia of Philosophy*, which offers a kind of Cliff Notes version overviewing various views on events[9] (among many other articles in philosophy). At least five or six strains of thought argue the nature of events. The fact we have no real intellectual consensus as to *What is an event?* after 2500 years suggests both that, it is a good question, but also that any answer is unlikely to find consensus. Nonetheless, the question of what is an event provides a good microcosm for understanding Peirce's worldview. For Peirce, "We perceive objects brought before us; but that which we especially experience — the kind of thing to which the word 'experience' is more particularly applied — is an event. We cannot accurately be said to perceive events." (1897, CP 1.336) He further states that "If I ask you what the actuality of an event consists in, you will tell me that it consists in its happening *then* and *there*. The specifications *then* and *there* involve all its relations to other existents. The actuality of the event seems to lie in its relations to the universe of existents." (1903, CP 1.24)

Though events are said to occur, to happen, or to take place, entities are said to exist. From Peirce again:

> "The event is the existential junction of *states* (that is, of that which in existence corresponds to a *statement* about a given subject in representation) whose combination in one subject would violate the logical law of contradiction. The event, therefore, considered as a junction, is not a subject and does not inhere in a subject. What is it, then? Its mode of being is existential quasi-existence, or that approach to existence where contraries can be united in one subject. Time is that diversity of existence whereby that which is existentially a subject is enabled to receive contrary determinations in existence." (1896, CP 1.494)

As Peirce says, "individual objects and single events cover all reality ...." (1905, CP 5.429)*

The event represents a juxtaposition of states, the comparison of the subject prior and after the event providing the basis for the nature of the event. Each change in state represents a new event, which can trigger new actions and reactions leading to still further events. Simple events represent relatively single changes in state, such as turning off a light switch or a bolt of lightning. More complicated events may involve multiple processes and potentially long durations, such as epoch, ages, or even geological eras. Havel insists that we should distinguish things and events only by differences in time scale: "In the world of all scales there is no essential difference: things are just long-lasting events and events are just short-lived things, where -long- and -short are relative with respect to our temporal scale perspective."[11]

How to characterize events provides a kind of Rorschach test for how one views reality. Events are like the spark that leads us to understand *actions* better and what emerges from them, which in turn helps us better understand predicates and relations. What we learn from Peirce is that *events* are quasi-entities, based on time rather than space, and, like entities, are a Secondness. Like entities, we can name events and intrinsically inspect their attributes. Events may also range from the simple to the triadic and durative.[12] Events are the first portions of activity and process cascades, and can stimulate such seemingly non-energetic actions like thought. Thought, itself, may be a source of further events and action, as may be the expressions of our thought, symbols. Actions always carry with them a reaction, which can itself be the impetus for the next action in the event cascade. Events are the real triggering and causative factors in reality. Entities are a result and manifestation of events. Events, like entities, are Secondness, or what we call *particulars*.

HIERARCHIES IN KNOWLEDGE REPRESENTATION

The human propensity to categorize is an attempt to make sense of the world. We base the act of categorization on how to group things and how to relate those things

* Here, Peirce uses a different sense for reality than his later belief that the universal categories are real. Also, there are many other useful statements by Peirce regarding events; see[10].

and groups to one another. Categorization demands that we characterize or describe the things of the world using what we have termed *attributes* to find similarities.* We may also categorize based on the relationships of things to external things.† No matter the method, the results of these categorizations are often hierarchical, reflective of what we see in the natural world. We see hierarchies in Nature based on bigger and more complex things built from simpler things, sometimes based on fractals or cellular automata or based on the evolutionary relationships of lifeforms. According to Annila and Kuismanen, "various evolutionary processes naturally emerge with hierarchical organization."[23] Hierarchy — and its intimate connection with categorization and categories — is thus fundamental to the why and how we can represent knowledge for computable means.

Depending on context, we can establish hierarchical relationships between types, classes or sets, with instances or individuals, with characteristics of those individuals, and between all of these things. The terminology differs by context, and sometimes the syntax may also carry a formal understanding of how we can process and compute these relationships. Nillson provides a general overview of these kinds of considerations with a useful set of references.[14]

Types of Hierarchical Relationships

As early as 1997 Doyle noted in the first comprehensive study of KR languages, "Hierarchy is an important concept. It allows economy of description, economy of storage and manipulation of descriptions, economy of recognition, efficient planning strategies, and modularity in design."[4] He also noted that "hierarchy forms the backbone in many existing representation languages."[15]

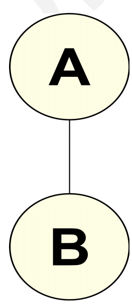

Figure 7-1: Direct Hierarchy

The basic idea of a hierarchy is that some item ('thing') is subsidiary to another item. Categorization, expressed both through the categories themselves and the process of how one splits and grows categories, is a constant theme in this book. The idea of hierarchy is central to a category or other such groupings and how we tie those cate-

* The most common analogous terms to attributes are *properties* or *characteristics; in* the OWL language used by KBpedia, we assign attributes to instances (called individuals) via property (relation) declarations.

† The act of categorization may thus involve intrinsic factors or external relationships, with the corresponding logics being either *intensional* or *extensional, as* discussed further in *Chapter 8*.

gories or groupings together. A hierarchical relationship is shown diagrammatically in *Figure 7-1* with A or B, the 'things,' shown as *nodes.* All this diagram is saying is that A has some form of superior or superordinate relationship to B (or *vice versa*, that B is subordinate to A). This hierarchical relationship is a direct one, but one of unknown character. Hierarchies can also relate more than two items:

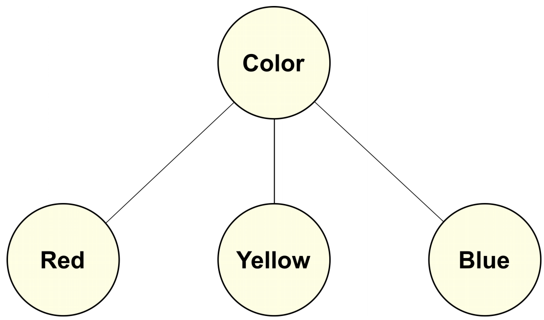

Figure 7-2: Simple Hierarchy

In this case, the labels of the items may seem to indicate the hierarchical relationship, but relying on labels is wrong. For example, let's take this relationship, where we show the mixed nature of primary and secondary colors:[16]

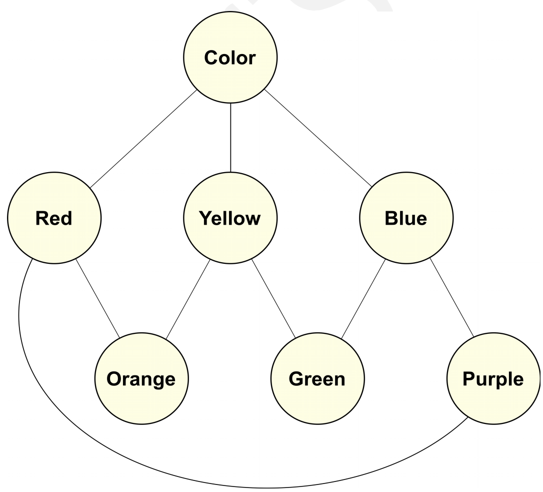

Figure 7-3: Multiple Intensional Hierarchy

Yet perhaps our intent was instead to provide a category for all colors lumped together, as instances of the concept 'color' show in *Figure 7-4* below.

The point is not to focus on colors – which are, apparently, more complicated to model than first blush – but to understand that hierarchical relations are of many types and what one chooses about a relation carries with it logical implications, the logic determined by the semantics of the representation language used and how we represent it.

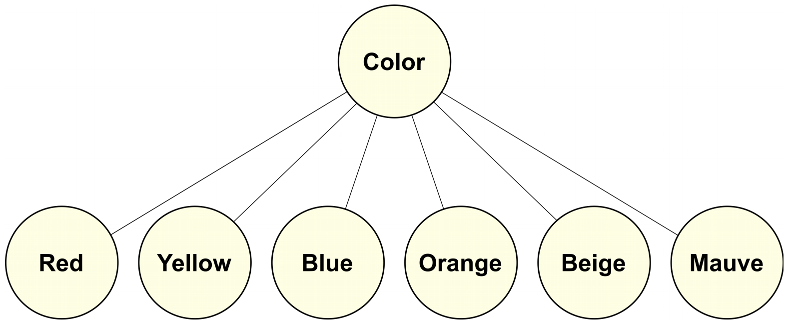

Figure 7-4: Extensional Hierarchy

Peirce's concept of *prescission* captures the most fundamental expression of a hierarchical relationship, stated as the relation, *prescind*.[17] Here are some quotes of how Peirce described the somewhat tricky method of *prescission*:

> "There are three distinct kinds of separation in thought. They correspond to the three categories. Separation of Firstness, or Primal Separation, called *Dissociation*, consists in imagining one of the two separands without the other. It may be complete or incomplete. Separation of Secondness, or Secundal Separation, called *Precission*, consists in supposing a state of things in which one element is present without the other, the one being logically possible without the other. Thus, we cannot imagine a sensuous quality without some degree of vividness. But we usually *suppose* that redness, as it is in red things, has no vividness; and it would certainly be impossible to demonstrate that everything red must have a degree of vividness. Separation of Thirdness, or Tertial Separation, called *discrimination*, consists in representing one of the two separands without representing the other." (1903, EP 2:270)

And,

> "But *prescission*, if accurately analyzed, will be found not to be an affair of attention. We cannot prescind, but can only distinguish, color from figure. But we can prescind the geometrical figure from color; and the operation consists in imagining it to be so illuminated that its hue cannot be made out (which we easily can imagine, by an exaggeration of the familiar experience of the indistinctness of hues in the dusk of twilight). In general, prescission is always accomplished by imagining ourselves in situations in which certain elements of fact cannot be ascertained. This is a different and more complicated operation than merely attending to one element and neglecting the rest." (1903, CP 2.248)

Prescission is an asymmetrical separation of two elements objectively considered; it is a logical operation that does not make any ontological or epistemological assumptions about the two elements being considered.[18] The process works by examining the two elements in isolation and then asking whether one might be possible or exist without the other. We can *dissociate* red from blue or a triangle from a square, but we can not *prescind* different colors or shapes. However, we *can* prescind color from shape because color is not possible without having a spatial aspect, but we *can not* prescind shape from color because a shape may exist without color. These distinctions are not grounded in experience, nor are they subjective, important to Peirce in finding a realistic logic. Prescission carries no connotation of meaning.

We can apply this same process to the ideas of generals and particulars. The general type of 'man' cannot be prescinded from a single, individual 'man' because we cannot conceive of a general type of 'man' without conceiving of individual 'men.' On the other hand, I can prescind the individual 'man' from the general type of 'man' since the idea of the general 'man' does not depend on the existence of any individual 'man.' Peirce uses the same process of prescission to prescind the concepts of First, Second and Third. Then, through a method he termed *hypostatic abstraction* (*e.g.*, CP 4.235), which is how to turn sign predicates into subjects (such as turning the predicate 'collect' into a general, singular of 'collection'), Peirce names the universal categories of Firstness, Secondness, and Thirdness.*

A	subsumes	B
A	prescinds	B
A	is more basic than	B
A	is a superClassOf	B
A	is more fundamental than	B
A	is broader than	B
A	includes	B
A	is more general	B
B	is-a	A
A	is parent of	B
A	has member	B
A	has an instance of	B
A	has attribute	B
A	has part	B

Table 7-2: Example Hierarchical Relationships

In use, we may see a variety of hierarchical relationships. *Table 7-2* shows some (vernacular) examples one might encounter. Again, though we have now labeled the relationships, which in a graph representation are the *edges* between the *nodes*, it is

* 'Prescind' is often more clearly stated as 'prescinded from.' Roughly equivalent phrases are to 'leave out of consideration,' 'separate from something,' or 'withdraw attention from.'

still unclear the populations to which these relations may apply and what their exact semantic relationships may be.

Table 7-3 shows some hierarchical relations that one might want to model, and whether the item resides in the universal categories of Firstness, Secondness or Thirdness:

Firstness		**Secondness**		**Thirdness**
attribute	—	token (instance)		
		sibling		
		|		
		sibling		
		child	—	parent
		|		
		parent		
		token	—	type
		part	—	whole
		|		
		whole		
		sub	—	super
				|
				sub

Table 7-3: Possible Pairwise (—) Hierarchical Relationships

Note that, depending on context, some of the items may reside in either Secondness or Thirdness (depending on whether the *referent* is a particular instance or a general). Also note the familial relationships shown: child-parent-grandparent and child-child relationships occur in actual families and as a way of talking about inheritance or relatedness relations. The idea of type or is-a is another prominent one in ontologies and knowledge graphs. Natural classes or kinds, for example, fall into the type-token relationship. Also note that mereological relationships, such as part-whole, may also leave open ambiguities. We also see specific pairs, such a sub-super, child-parent, or part-whole, need context to resolve the universal category relation.

Reliance on item labels alone for the edges and nodes, even for something as seemingly straightforward as color or pairwise relationships, does not give us sufficient information to determine how to evaluate the relationship nor how to organize properly. We thus see in knowledge representation that we need to express our relationships explicitly. Labels are merely assigned names that, alone, do not specify the logic mode, what populations are affected, or even the exact nature of the relationship. Without these basics, our knowledge graphs cannot be computable. Well over 95% of the assignments in contemporary knowledge bases have this item-item character. We need interpretable relationships to describe the things that populate our domains of inquiry to categorize that world into bite-sized chunks.

Salthe categorizes hierarchies into two types: compositional hierarchies and subsumption hierarchies.[16] [17] A subsumptive hierarchy 'subsumes' its children, and a compositional hierarchy is 'composed' of its children. Mereological and part-whole hierarchies are compositional, as are entity-attribute. Subsumption hierarchies are ones of broader than, familial, or evolutionary. Cottam *et al.* believe hierarchies are so basic as to propose a model abstraction over all hierarchical types, including levels of abstraction.[18]

These discussions of structure and organization are helpful to understand the epistemological bases underlying various kinds of hierarchy. We should also not neglect recursive hierarchies, such as fractals or cellular automata, which are also simple, repeated structures commonly found in nature. Fortunately, Peirce's universal categories provide a powerful and consistent basis for us to characterize these variations, and his notion of prescission also helps adjudicate logical hierarchical relationships. When paired with logic and the KR languages discussed in *Chapter 8*, and with "cutting Nature at its joints" in *Chapter 5*, we end up with an expressive grammar for capturing all kinds of internal and external relations to other things.

So far we have learned that most relationships in contemporary knowledge bases are of a noun-noun or noun-adjective nature, which I have loosely lumped together as hierarchical relationships. These relationships span from attributes to instances (individuals) and classes[21] or types, with and between one another. We have learned Peirce's logical concept of how to prescind a superordinate concept from a subordinate one. We have further seen that labels either for the subjects (nodes) or their relationships (edges) are an insufficient basis for computers (or us!) to reason over. Mostly, we have come to see that we need to ground our relationships in specific semantics and logics for reasoning machines to process our representations without ambiguity.

Structures Arising from Hierarchies

Structure is a tangible part of thinking about a new KR installation since we may apply many analytic choices against the knowledge artifact. Different kinds of structure are best for various tools or kinds of analysis. The types of relations chosen for the artifact affects its structural aspects. These structures can be as small and straightforward as a few members in a list, to the entire *knowledge graph* fully linked to its internal and external knowledge sources. Knowledge structures arise from the various hierarchical relationships just discussed. Here are some of the prominent types of structures that may arise from connectedness and characterization hierarchies:

- *Lists* — unordered members or instances, with or without gaps or duplicates, useful for bulk assignment purposes. Lists occur through a direct relation assignment (*e.g.*, `rdf:Bag`);
- *Neural networks* (*graphs*) — graph designs based on connections modeled on biological neurons, still in the earliest stages for relations and KR formalisms;[22]
- *Ontologies* (*knowledge graphs*) — sometimes ontologies are treated as synonymous with knowledge graphs, but more often as a superset that may allow more control and semantic representation.* Ontologies are a central design feature of KBpedia;[23]
- *Parts-of-speech* — a properly designed ontology has the potential to organize the vocabulary of the KR language itself into corresponding parts-of-speech, which aids some methods of natural language processing;
- *Sequences* — ordered members or instances, with or without gaps or duplicates, useful for bulk assignment purposes. Sequences occur through a direct relation assignment (*e.g.*, `rdf:Seq`);
- *Taxonomies* (*trees*)— trees are subsumption hierarchies with single or multiple class inheritance for instances; most knowledge graphs allow multiple inheritances; or
- *Typologies* — are essentially multi-inheritance taxonomies, with the hierarchical organization of types as natural as possible. Natural types (classes or kinds) enable us to make the largest number of disjoint assertions, leading to efficient processing and modular design. Typologies are a central design feature of KBpedia; see *Chapter 10*.

Typically KR formalisms and their internal ontologies (taxonomy or graph structures) have a starting node or *root*, often called 'thing,' 'entity' or the like. Close inspection of the choice of the root may offer important insights into different KR language philosophies. 'Entity' as a root, for example, is not compatible with a Peircean interpretation, since all entities are within Secondness, one of the three subsidiary branches in our main KR structure.

KBpedia's foundational structure is the subsumption hierarchy shown in the KBpedia Knowledge Ontology (KKO) — that is, KBpedia's upper ontology — and its nodes derive from the universal categories. The terminal, or *leaf*, nodes in KKO each tie into typologies. *Types* are the constituents of a typology. Types, which are generals along with typologies, are the classification of natural kinds of instances as determined by shared attributes (though not necessarily the same values for those attributes). Most of the types in KBpedia are composed of entities, but attributes and relations also have aggregations of types. In turn, a *typology* is a hierarchical classification of related types as determined by the essence or characteristics of its root. Subsequent chapters discuss these items in some detail; *Appendix B* describes KKO.

* RDF graphs are more akin to the first sense; OWL 2 graphs more to the latter; see next chapter.

Of course, choice of a KR formalism and what structures it allows must serve many purposes. We desire uses of the KR formalism and the knowledge graph to include knowledge extension and maintenance, record design, querying, reasoning, graph analysis, logic and consistency tests, planning, hypothesis generation, question and answering, and subset selections for external analysis. We are often best supporting other tasks such as machine learning, natural language processing, data wrangling, statistical and probabilistic analysis, search indexes, and other data- and algorithm-intensive applications using dedicated external applications. We have as a goal to build structures into the KR installation to support these kinds of uses, or to export data suitable to external applications. *Chapter 12* expands on these platform considerations.

A THREE-RELATIONS MODEL

If hierarchy provides the basis for the scaffolding in a knowledge graph, then *actions* offer the means to make a knowledge system dynamic. Moving beyond static knowledge representations is the way for these systems to support active learning, respond to sensors, plan, hypothesize, and solve problems. Peirce's universal categories and these hierarchical perspectives dovetail nicely into a three-relations model that captures all aspects of knowledge representation to support the full slate of anticipated artificial intelligence applications. Relations are the way we describe connections among things, including attributes which we only express for subjects.

Guarino, in some of the earliest (1992) writings leading to semantic technologies, had posited knowledge bases split into concepts, attributes, and relations.[24] This split was close to my thinking and provided comfort since it arose in the earliest days of the semantic Web.[25] Some of the impressive work by Sekine[26] extending the concept of entity types influenced me greatly. Still, I was confused by the mixing of attributes and entities; indeed, most practitioners do not appreciate or employ the purposeful separation of attributes from other relations, let alone entities. It was only after study of Peirce that I realized he had a way to untangle the knot of attributes, events, relations, actions, perceptions, thoughts, and belief. His '*architectonic*' began providing answers to epistemological questions across the board. It still does.

Besides Peirce, I studied thinkers across history who may have tackled fundamental concepts in knowledge organization. Aristotle's categories were influential, and have mostly stood the test of time and figured prominently in my thinking, as they did for Peirce. Peirce was a student of Kant and Hegel (as well, in contrast, Descartes), and the logicians DeMorgan, Boole and Venn, but he created a unique synthesis. I also reviewed efforts such as Sarbo's to apply Peirce to knowledge bases,[27] as well as most other approaches discussing Peirce with some correspondence to KBs.[28] [29] [30]

Our resulting three-relations model is consistent with Peirce's thinking, even though he never had today's concepts of digital knowledge representation as an objective. For example, he labeled one of his sections "The Conceptions of Quality, Relation and Representation, Applied to this Subject" ("Upon Logical Comprehension and Extension"; 1867, CP 2.418). Thirty-five years later, Peirce still held to this split,"... there are but three elementary forms of predication or signification, which as I originally named them (but with bracketed additions now made to render the terms more intelligible) were qualities (of feeling), (dyadic) relations, and (predications of) representations." (1903, EP 2:424; CP 1.561)

Thirdness is the sauce that gives meaning to what is different in Peirce's architectonic over standard knowledge representations. Too many pivotal problems we cannot address with dichotomous worldviews. Disambiguation is made difficult without context. The world is probabilistic. Chance happens. New information is a barrage, we continuously seek knowledge, and beliefs evolve and change. Though we may partially describe context with nouns related to perception, situations, states, and roles, we ultimately require an understanding of events, actions, and relations. Until these latter factors are better captured and understood, our ability to establish context remains limited. Peirce elaborates:

> "Now every simple idea is composed of one of three classes; and a compound idea is in most cases predominantly of one of those classes. Namely, it may, in the first place, be a quality of feeling, which is positively such as it is, and is indescribable; which attaches to one object regardless of every other; and which is *sui generis* and incapable, in its own being, of comparison with any other feeling [a*ttributes*], because in comparisons it is representations of feelings and not the very feelings themselves that are compared. Or, in the second place, the idea may be that of a single happening or fact, which is attached at once to two objects, as an experience, for example, is attached to the experiencer and to the object experienced [*external relations*]. Or, in the third place, it is the idea of a sign or communication conveyed by one person to another (or to himself at a later time) in regard to a certain object well known to both [r*epresentations*]." (1905, CP 5.7) (labeling brackets added.)

We now have the basis to define the three modes of relations within KBpedia. The first of these is the grouping of a*ttributes*, the relationship of a subject with its intrinsic qualities or characteristics, which are a Firstness within Peirce's universal categories. The second of these modes is *external relations*, which are all of the ways a *particular* or *general* may relate to another *particular* or *general*. These include hierarchical relations other than attributes (which are monadic).* Relations of a*ction* (one thing affecting another) or *perception* (one thing experiencing an external change) are external relations, which are a Secondness within the universal categories. The third mode of relations we call r*epresentations* since these are the ways we describe, point to, or otherwise indicate the thing at hand. These relations give our subjects perspective and meaning, though we cannot easily reason over these relations. They are a Thirdness within Peirce's universal categories. These constructs are central to our approach to knowledge representation.

* However, we can type attributes, so it is possible to organize and reason over them.

Attributes, the Firstness of Relations

Attributes are the intensional characteristics of an object, event, entity, type (when viewed as an instance), or concept. The relationship is between the individual instance (or particular) and its attributes and characteristics, in the form of *A:A*. Attributes may be intrinsic characteristics or essences of single particulars, such as colors, shapes, sizes, or other descriptive characteristics. Attributes may be adjunctual or accidental happenings to the particular, such as birth or death. Attributes may be contextual for placing the particular within time or space or external circumstances, absent having a direct relationship (in that case it is an *external relation*).

Attributes are specific to the individual, and only include events that are notable for the individual. They are a Firstness, and in totality try to capture the complete characteristics of the individual particular, which is a Secondness. Since attributes are the properties of an entity, we can better interoperate entity data by concentrating on those aspects that let us match data in one set of records to similar data in different records. In the next chapter, we will discuss building a new vocabulary and structure upon RDF to provide more sophisticated handling of 'properties' than RDF or OWL alone can offer in their native forms.

Calling out attributes for such attention is not new. The attributes-relation split has not been an uncommon one in the KB literature,[28] [32] though it is not accepted canon and is infrequent in other knowledge representations. Philosophers draw distinctions about intrinsic *v* extrinsic properties[33] or intensionality *v* extensionality.[34] For conceptual models with specific reference to ontologies, Wand *et al.*[35] in 1999 were making the distinction between *intrinsic properties* (akin to what we term attributes herein) and *mutual properties* between things (what we term external relations). Unfortunately, at that time, the conventions of RDF had not yet become prevalent, and the idea of annotation properties had not yet emerged (from OWL). These later distinctions are important, but the Wand *et al.* discussion still is helpful to elucidate the same pragmatic and theoretical considerations.

With all of this discussion of attributes the attentive reader might be confused: Are attributes not nouns or adjectives that seem similar to objects as we discussed for hierarchies? Alternatively, are attributes a more verb-like relation? The answer, naturally, is that it depends. When we think about an attribute as some quality of something, we reify it as a noun and make it its object. Considered in this manner, the 'idea' of an attribute makes it a real thing, and a Secondness in that reified state (which, of course, is not the same as the underlying thing). Without that relation to the something, it does not exist, which makes it only an ephemeral quality, a Firstness. We can both describe and relate attributes, depending on context. It is this kind of contextual lens that makes Peirce's universal categories so powerful.

External Relations, the Secondness of Relations

External relations are actions or assertions between an event, entity, type, or concept and another particular or general. An external relationship has the form of *A:B*.

External relations may be simple ones of a direct relationship between two different instances. External relations may be copulative by combining objects or asserting membership, quantity, action or circumstance. External relations may be mediative to provide meaning, context, relevance, generalizations, or other explanations of the subject vis-à-vis the external world. External relations are extensional.

All actions are external relations. Actions may be reactions to perceptions or stimuli. Actions may be energetic, arising from the subject and affecting the external environment in some way. We may understand some actions as a basis of thought, which results in new actions or modified concepts or thought. External relations are by definition a Secondness. Notice how these three groupings of external relations are themselves an example of the universal categories. It is in this manner that bigger, more abstract notions may be broken down into more manageable pieces by employing the universal categories.

Representations, the Thirdness of Relations

The third category in our model of relations is the least used and, perhaps, the most confused regarding how other KR systems treat their scope. *Representations* are signs (1905, CP 8.191) and the means by which we point to, draw attention to, or designate, denote or describe a particular object, entity, event, type or general. A representational relationship has the form of *re:A* (about *A*). Representations can be designative of the subject, that is, be icons or symbols (including images, labels, definitions, and descriptions). Representations may be indexes that more-or-less help situate or provide a traceable reference to the subject. Representations may be associations, resemblances, and likelihoods about the subject, more often of indeterminate character (such as a probability assignment). Representations are the Thirdness of relations. Representations cannot be deductively reasoned over, but some characteristics may be derived or analyzed through inductive or abductive inferential means.

For example, *annotations* are representations. Annotations capture the circumstances or conditions or contexts or observations for the thing at hand. Where did we discover or find it? When did we find or elaborate on it? By whom or when was it found or elaborated? What is our commentary about it? While these are all external elaborations of the thing at hand, and not intrinsic to the nature of the thing, they are all characterizations about a given thing. In these regards, annotations have as their focus a given object, similar to what is valid for attributes. We cannot deductively reason over annotations, though annotations play pivotal roles. Annotations are an essential means for tagging, matching and slicing-and-dicing the information space. *Metadata* is a similar concept, more oriented to provenance and description.

Labels, which are also representations as are definitions, are the means to broaden the correspondence of real-world reference to match the true referents or objects in the knowledge base.* Broader reference enables us not to limit referents to any given label or string. In best practice, labels should reflect all of the various ways a given object may be identified (synonyms, acronyms, slang, jargon, all by language

* See the discussion of *semsets* in *Chapter 10*.

type). These considerations improve the means for tagging, matching, and slicing-and-dicing, even if we can not reason over the annotations.

The Basic Statement

We now have a starting grammar to talk about the things of the world, and the relations that place them into context with the external world. We have our subjects and objects (*nodes*) and our model of how to relate them (*edges*). The combination of these parts gives us the basis for making basic statements about the world, what we assert as statements of *fact*.[*] Practitioners call this primary construct a *triple*. Triples are statements in the RDF language that relate a *subject* and *object* through a connecting *property* (or *predicate*). Triples take the form of *s – p - o*, with the subject and property (and object optionally) referenced by an *IRI* (Web link). I expand on the construct of triples in the next chapter; see also the discussion related to *Figure 1-2*.

A proposition captures a relation, an assertion about the subject. “Any portion of a proposition expressing ideas but requiring something to be attached to it in order to complete the sense, is in a general way relational. But it is only a *relative* in case the attachment of indexical signs will suffice to make it a proposition, or, at least, a complete general name.” (1897, CP 3.463) “But the Logic of Relations has now reduced logic to order, and it is seen that a proposition may have any number of subjects but can have but one predicate which is invariably general.” (1903, CP 5.151)

We now have a much clearer way for how to build up the assertions in our knowledge representations. We now know that attributes are a Firstness in the universal categories, that Secondness captures all events, entities, and relations, and that Thirdness provides the context, meaning, and ways to indicate what we refer to in the world. We see how context is operative: relations as a construct, for example, are in Secondness, but within relations the mode of representations is in Thirdness. We now have a framework of triadic relations in attributes, external relations and representations for how to describe things and relate them to one another. Peirce and his architectonic provides the richest, most expressive basis for capturing human language and conducting logical reasoning, both for individuals (particulars) and concepts (generals). This starting grammar sets the foundation for us to compute and reason over human language for modern KR purposes.[†]

Chapter Notes

1. Some material in this chapter was drawn from the author’s prior articles at the *AI3:::Adaptive Information* blog: “Conceptual and Practical Distinctions in the Attributes Ontology” (Mar. 2015); "KBpedia Relations, Part I: Smarter Knowledge Graphs” (May 2017); “KBpedia Relations, Part II: An Event-Action Model” (May 2017); “KBpedia Relations, Part III: A Three-Relations Model” (May 2017).

* If validated, they are indeed *fact* assertions. However, as discussed elsewhere, facts are subject to question and have some degree of *fallibility*; acceptance of an assertion as fact is a matter of *belief*

† Additional Peirce quotes may be found in my initial article.[36]

2. Isnenghi, A., "A Semiótica de CS Peirce e a Gramática Especulativa de Modistae (or, "C.S. Peirce's Semiotic And Modistae's Grammatica Speculativa")," *Cognitio-Estudos: Revista Eletrônica de Filosofia*, vol. 1809, 2008.
3. Peirce, C. S., "Description of a Notation for the Logic of Relatives, Resulting from an Amplification of the Conceptions of Boole's Calculus of Logic," *Memoirs of the American Academy of Arts and Sciences*, vol. 9, 1870, pp. 317–378.
4. In OWL 2, "thing" is the *root* node, with all other vocabulary items being subsidiary to it. Here, we are using "thing" more in keeping with the common vernacular.
5. The role for the label "entity" can also refer to what is known as the *root* node in some systems such as SUMO. In the OWL language and RDF data model we use, we know the root node as "thing.". Our use of the term "entity" is much different from SUMO and resides at a subsidiary place in the overall TBox hierarchy (see *Chapter 8*). In this case, and frankly for most semantic matches, equivalences should be judged with care, with context the crucial deciding factor.
6. Chen, P. P.-S., "The Entity-Relationship Model—Toward a Unified View of Data," *ACM Transactions on Database Systems (TODS)*, vol. 1, 1976, pp. 9–36.
7. They are, however, *real*, since their type concept exists independent of our thinking about it. Peirce's insistence that generals may be real helps to situate him among a common split of philosophers into the nominalist, idealist, and realist camps.
8. See https://en.wikipedia.org/wiki/Topic_map.
9. Casati, R., and Varzi, A., "Events," *The Stanford Encyclopedia of Philosophy* Available: https://plato.stanford.edu/archives/win2015/entries/events/.
10. Bergman, M. K., "KBpedia Relations, Part II: An Event-Action Model," *AI3:::Adaptive Information*, May 2017.
11. Havel, I. M., "Scale Dimensions in Nature," *International Journal Of General Systems*, vol. 24, 1996, pp. 295–324.
12. Within events, we can also categorize according to the three universal categories. The unexpected flash or shock is a Firstness within events. Peirce's doctrine of tychism places a central emphasis on chance, which he views as the source of processes in nature such as evolution and the "surprising fact" that causes us to re-investigate our assumptions leading to new knowledge." We more commonly associate an event with action, and that is indeed a major cause of events. (Though, chance events or accidents, as an indeterminate group, may trigger events.) An action is a Secondness, however, because it is always paired with a reaction. Reactions may then cause new actions, itself a new event. In this manner activities and processes can come into being, which while combinatorial and compound, can also be called events, including those of longer duration. That entire progression of multiple actions represents increasing order, and thus the transition to Thirdness. Peirce makes the interesting insight that thoughts are events, too. "Now the logical comprehension of a thought is usually said to consist of the thoughts contained in it; but thoughts are events, acts of the mind. Two thoughts are two events separated in time, and one cannot literally be contained in the other." (1868, CP 5.288).
13. Annila, A., and Kuismanen, E., "Natural Hierarchy Emerges from Energy Dispersal," *Biosystems*, vol. 95, Mar. 2009, pp. 227–233.
14. Nilsson, J. F., "Ontological Constitutions for Classes and Properties," *Conceptual Structures: Inspiration and Application*, H. Schärfe, P. Hitzler, and P. Øhrstrøm, eds., Aalborg, Denmark: Springer, 2006, pp. 35–53.
15. Doyle, J., *Hierarchy in Knowledge Representations*, MIT Artificial Intelligence Laboratory, 1977.
16. J W von Goethe (1749-1832) first explicated the standard 3-color scheme. What is more commonly used in design is a 4-color scheme from Ewald Hering (1834-1918).
17. Peirce also termed this concept *precission*, *prescisive abstraction*, *prescision*, or *precisive abstraction* (1902, CP 4.235). It comes from the same root as precision in measurements but has a different meaning as described in the text, which is one reason we prefer the spelling of *prescission* to distinguish it as much as possible.
18. Houser, N., "Peirce, Phenomenology, and Semiotics," *The Routledge Companion to Semiotics*, P. Cobley, ed., London ; New York: Routledge, 2010, pp. 89–100.
19. Salthe, S., "Hierarchical Structures," *Axiomathes*, vol. 22, Sep. 2012.
20. There is a very helpful 25-page listing of references dealing with "hierarchy" at the conclusion of Salthe's 2012 paper, Hierarchical Structures.

21. In the OWL 2 language used by KBpedia, a *class* is any arbitrary collection of objects. A class may contain any number of *instances* (called *individuals*), or a class may be a subclass of another. Instances and subclasses may belong to none, one or more classes. Both extension and intension may be used to assign instances to classes.
22. Santoro, A., Raposo, D., Barrett, D. G. T., Malinowski, M., Pascanu, R., Battaglia, P., and Lillicrap, T., "A Simple Neural Network Module for Relational Reasoning," *arXiv:1706.01427 [cs]*, Jun. 2017.
23. In the semantic Web space, "ontology" was the original term because of the interest to capture the nature or being (Greek ὄντως, or ontós) of the knowledge domain at hand. Because the word 'ontology' is a bit intimidating, a better variant has proven to be the knowledge graph (because all semantic ontologies take the structural form of a graph). In this book, I tend to use the terms ontology and knowledge graph interchangeably.
24. Guarino, N., "Concepts, Attributes and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases.," *Data & Knowledge Engineering*, vol. 8, 1992, pp. 249–261.
25. I raise the early work by Guarino for a reason. We, the community of KR practitioners, have not gotten our basic grammar right about how we think about these problems. Most everyone still gets bollixed up trying to handle concepts like relations (for me, split into the three categories of attributes, external relations, and representations), events, generals (types or classes), and particulars (individuals or instances). Peircean principals give us logical and defensible ways to think about these problems. That approach strikes me as superior to heated assertions that often lack logical underpinnings.
26. Sekine, S., "Extended Named Entity Ontology with Attribute Information," *Proceedings of the Sixth International Language*, 2008, pp. 52–57.
27. Breemen, A. J. J. V., and Sarbo, J. J., "The Machine in the Ghost: The Syntax of Mind," *Signs-International Journal of Semiotics*, vol. 3, 2009, pp. 135–184.
28. Lehmann, F., and Wille, R., *A Triadic Approach to Formal Concept Analysis*, Heidelberg: Springer Berlin, 1995.
29. Farkas, J. I., "A Semiotically Oriented Cognitive Model of Knowledge Representation," [Sl: sn], 2008.
30. Sowa, J. F., "Top-level Ontological Categories," *International Journal of Human-computer Studies*, vol. 43, 1995, pp. 669–685.
31. Guarino, N., "Some Organizing Principles for a Unified Top-Level Ontology," *AAAI Spring Symposium on Ontological Engineering*, 1997, pp. 57–63.
32. Lin, Y., Liu, Z., and Sun, M., "Knowledge Representation Learning with Entities, Attributes and Relations," *Ethnicity*, vol. 1, 2016, pp. 41–52.
33. Weatherson, B., and Marshall, D., "Intrinsic vs. Extrinsic Properties," *The Stanford Encyclopedia of Philosophy* Available: https://plato.stanford.edu/archives/fall2017/entries/intrinsic-extrinsic/.
34. At least for Carnap, he thought "...the full meaning of a concept is constituted by two aspects, its intension and its extension. The first part comprises the embedding of a concept in the world of concepts as a whole, i.e., the totality of all relations to other concepts. The second part establishes the referential meaning of the concept, i.e., its counterpart in the real or in a possible world.".
35. Wand, Y., Storey, V. C., and Weber, R., "An Ontological Analysis of the Relationship Construct in Conceptual Modeling," *ACM Transactions on Database Systems (TODS)*, vol. 24, Dec. 1999, pp. 494–528.
36. Bergman, M. K., "KBpedia Relations, Part III: A Three-Relations Model," *AI3:::Adaptive Information*, May 2017.