

Ontologies across disciplines

Matthias Nickles, Adam Pease, Andrea C. Schalley, and Dietmar Zaefferer

The English word *ontology*, together with its counterparts in many languages, has made a breathtaking career during the last decades especially in information science, but also in other disciplines. Since its definitions vary considerably within and especially across disciplines, and since this volume, although clearly focused on linguistic matters, is conceived as tying together several disciplines, it seems appropriate to provide a short survey of these uses in order to make the different contributions and their interconnections more accessible for those readers who are not familiar with all the fields that are represented between the covers of this book (presumably the majority).

1. Readings of *ontology*: Carving up a conceptual space

1.1. The major dimensions of variation

The traditional notion of ontology has a long and venerable history in philosophy. The most usual word for it, however, the compound built from the Greek forms *onto-* ‘of being’ and *logos* ‘speech, reason’ in the guise of its later derivative *logia* ‘science’, is a comparatively recent invention. It seems to have originated in the context of the early Enlightenment, since its first attested appearance in print is in 1606 on the front page of the textbook *Ogdoas Scholastica* (‘Scholastic Eightfold’) by Jacob Lorhard,¹ and it became popular about one century later when Christian Wolff used it in the title of his 1729 book *Philosophia Prima sive Ontologia* (First Philosophy or Ontology). There, Wolff gives the following definition: *Ontologia seu Philosophia Prima est scientia entis in genere, seu quatenus ens est* (Ontology or First Philosophy is the science of Being in general or as Being).

But the study of Being as Being goes back at least to Aristotle’s *Metaphysics*. To quote the philosopher Nino Cocchiarella:

Aristotle was the founder not only of logic in western philosophy, but of ontology as well, which he described in his *Metaphysics* and the *Categories* as

a study of the common properties of all entities, and of the categorial aspects into which they can be analyzed. The principal method of ontology has been one or another form of categorial analysis, depending on whether the analysis was directed upon the structure of reality, as in Aristotle's case, or upon the structure of thought and reason, as, e.g., in Kant's *Critique of Pure Reason*. (Cocchiarella 2001: 117)

Despite its conciseness this characterization already makes it possible to distinguish two dimensions of variation along which notions of ontology vary. One is opened up by the common properties of all entities on the one hand (being as being) and their categorial aspects on the other (categories and kinds of entities), let us call this the *dimension of generality* (as opposed to specificity). The other one is opened up by the distinction between external reality and the contents of thought and reason. Let us call this *the dimension of objectivity* (as opposed to subjectivity). Since the two are orthogonal, we can imagine them as spanning a vertical plane where generality extends from its maximum at the top through increasing degrees of specificity to the lower bound of generality at the bottom, and where objectivity extends in the depth with its maximum at the foreground and increasing degrees of subjectivity towards the back (cf. Figure 1).

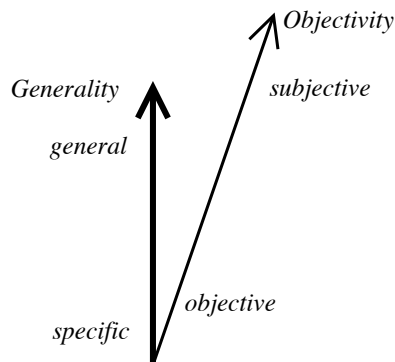


Figure 1. The generality-objectivity plane.

So when the *Dictionary of Philosophical Terms and Names* defines ontology as “Branch of metaphysics concerned with identifying, in the most general terms, the kinds of things that actually exist”, (<http://www.philosophypages.com/dy/o.htm#onty>) we can say that this notion of ontology places it close to the upper foreground in our picture. Sim-

ilarly when ontology is variously described as being concerned with the ultimate furniture, or the basic furniture or simply the furniture of the world, this corresponds to increasingly large regions on our plane from the top to the bottom. And when Aristotle is concerned more with the structure of reality itself, this notion of ontology is located in the foreground of our plane, whereas Berkeley's ([1710] 1999: §6) idealist view that *All the choir of heaven and furniture of the earth ... have not any subsistence without a mind* has its place considerably further back.

But how does this relate to Gruber's often-quoted definition of ontology for the purposes of Artificial Intelligence² as "an explicit specification of a conceptualization," where a "conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose" (Gruber 1993: 199)? Is there a connection at all or is this a case of homonymy, an entirely unrelated different use of the same orthographical form? Gruber adds: "The term is borrowed from philosophy, where an ontology is a systematic account of Existence." This is of course neither the notion of ontology we have just discussed (science of being as being; study of the common properties of all entities; branch of metaphysics concerned with the kinds of things that exist) nor something completely different, instead it is something closely related: The concept coded by the mass noun (no article, no plural) is that of a field of investigation, of a discipline ('science', 'study', 'branch'), the concept expressed by the count noun (taking articles and plural form) is that of a specific outcome of that kind of investigation ('account'): Aristotle's ontology, e.g., is different from Kant's, but both contribute to (the field of) Ontology. In order to visualize the distinction we will use an uppercase initial for the name of the field (domain of issues, etc.) – i.e. 'Ontology' – and a lowercase initial for the different views that are produced in the field – i.e. 'ontology'.

This kind of field-product polysemy is familiar from linguistics: *Syntax* as a mass noun means a field, a certain branch of linguistics; its different outcomes – like say Haider's syntax of German (Haider 1993) – are coded by the corresponding count noun. In fact, in linguistics there is a third use of the term syntax (and a second use of the count noun), one that relates to the subject matter of the second and first use, i.e., that subsystem of a language that constrains the building of phrases from word forms. So there is an object-level use of this term (syntax as language subsystem), a meta-level use (syntax as theoretic account of this subsystem), and in a sense a transmeta-level use (syntax as subfield or branch of linguistics). Is there a similar three-level distinction with *ontology*?

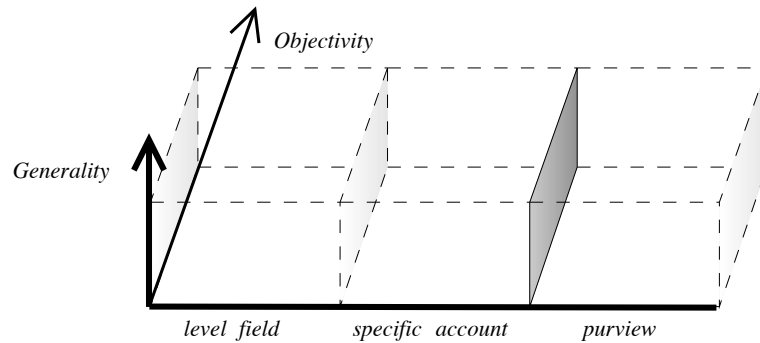


Figure 2. Conceptual space of the notion *ontology*

The answer is ‘yes’ and ‘no’. ‘Yes’ in the sense that of course there are three levels as well: What there is and its categories (object-level), specific accounts thereof (meta-level), and a field concerned with this subject matter and therefore with the production and discussion of specific accounts of being (yet another level). ‘No’ in the sense that only the last two levels are properly called ontology, the second one by transparent metonymic extension (and count noun formation) from the name for the third one, whereas the first one requires different means of expression³ such as *the real world* (as opposed to possible counterparts) or simply *reality* or rather its (ultimate or basic) furniture.

Whether we include the object-level in this overview or not (and in fact we should include it for the sake of being systematic), we now have introduced a third dimension of variation and thus created a conceptual space where variants and relatives of the notion ontology can be localized (cf. Figure 2): The vertical dimension reflects generality with the most general matters at the top; the depth dimension reflects objectivity with the most objectivist view at the front; and the horizontal dimension has three segments with the world and its aspects and parts at the right, the different accounts of it in the middle and the field(s) of Ontology at the left.

Note that the horizontal relations between the three blocks are somewhat heterogeneous because the field of Ontology is not located on a meta-level with respect to the different ontologies in the same way the latter are on a meta-level with respect to what they account for. That would be the level of a metaontology, an ontological account of different object ontologies. The field of Ontology is rather something that is concerned with being and its kinds by

producing and discussing ontologies. Obviously, the present book is a product of this (ever evolving) field, and what the present subsection tries to outline is a clarification of different concepts that have come to be called ‘ontology’.

It is now time to come back to the second dimension of variation, objectivity. The AI-reading of the notion of ontology mentioned above seems to be rather clear in this respect: “An *ontology* is an explicit specification of a conceptualization.” (Gruber 1993: 199) This points to a subjective notion of ontology: Not reality *per se* is its object, but reality under a given conceptualization. But the latter notion is explained by Gruber as follows: “conceptualization: the objects, concepts, and other entities that are presumed to exist in some area of interest and the relationships that hold [among] them” (Gruber 1993: 199). This is, to say the least, a little confusing: If your current area of interest is your desktop and if you presume that an object exists there, namely a pile of unread papers, is this object then part of a conceptualization? Certainly not. Gruber must mean a conceptualization of this object. But then what are the concepts that you presume to exist on your desktop? Probably there aren’t any, because wherever concepts exist, in Plato’s heaven or in people’s minds, they certainly do not exist on desktops. On the other hand, concepts fit much better in a conceptualization than objects. So the quoted definition is not very helpful, and we wouldn’t have bothered mentioning it at all if this kind of confusion of object- and meta-level didn’t seem to be quite widespread in the field.⁴

What Gruber must mean by conceptualization of a given domain is a system of concepts that adequately characterize everything that exists in that domain: individual concepts for individuals, property concepts for properties, relation concepts for relations, second order concepts for first order concepts, etc. But it is an open question whether these concepts are meant to be objective and the characterizations they provide are thought of as realistic (located at the foreground of our conceptual space), or more to the back in the sense of representing some other view. And maybe this is good as it is, because the depth dimension of the conceptual space of ontologies is the most difficult and philosophically most demanding one. It is to do with the independence or interdependence of Ontology and Epistemology, with questions of realism and opposing views, in short with the “most hotly debated issues in contemporary metaphysics” according to philosopher Alexander Miller (2005).

Then what is realism? “Realism is the thesis that the objects, properties and relations the world contains exist independently of our thoughts about them or our perceptions of them. Anti-realists either doubt or deny the exist-

tence of the entities the realist believes in or else doubt or deny their independence from our conceptions of them.” (Khleutzos 2004) Realism is rarely held across the board, philosophers rather tend to be realist about one domain and non-realist about another. John Searle, for instance, in his recent paper *What is an institution?* says:

[I]t is essential to distinguish between those features of the world that are totally independent of human feelings and attitudes, observer independent features, and those features of the world that exist only relative to human attitudes. . . . It is important to see that one and the same entity can have both observer independent features and observer dependent features, where the observer dependent features depend on the attitudes of the people involved. (Searle 2005: 3–4)

Searle’s aim in the paper just quoted from is “to explain how the ontology of institutions fits into the more basic ontology of physics and chemistry” (Searle 2005: 1) and the explanation he offers is the following:

[O]ne and the same phenomenon (object, organism, event, etc.) can satisfy descriptions under which it is non-institutional (a piece of paper, a human being, a series of movements) and descriptions under which it is institutional (a twenty dollar bill, the president of the United States, a football game). An object or other phenomenon is part of an institutional fact, *under a certain description of that object or phenomenon*. (Searle 2005: 12)

For Searle an institutional fact is something that has been collectively assigned a status function, and since this collective assignment presupposes some representation of it as having this function, institutions require (at least some primitive form of) language.

Although Searle devotes one section of his paper to *Language as the Fundamental Social Institution*, he is mainly concerned there with showing that language is a prerequisite for social institutions and not with discussing the status of language as a social institution itself. The reason may be this: If language were a social institution just as the others it would have to be a prerequisite of itself. The way out of this seeming circle is not hard to find: Each higher form of language requires only some more primitive form of language and so there is space for the evolution of language out of more primitive forms of representation sharing. Therefore an extension of Searle’s (or some similar) concept of an institution to include language as the fundamental in-

stitution along these lines is consistent and will be taken for granted in the following sections.⁵

Coming back to the definition of realism above (“the thesis that the objects, properties and relations the world contains exist independently of our thoughts about them or our perceptions of them”) we now see that the devil is in a certain detail, namely the reference of the possessive pronoun form *our*. Here we have two options. If we read *our* to include any rational subject, then we are forced to assume an at least partially non-realist position if we assume (as we probably should) that the world contains among other things institutions and institutional facts, since we have just subscribed to the view that their existence depends on thoughts about and perceptions of them. If we read it to include only the persons who are involved in the current reflection process, then we can maintain a completely realist position which has the interesting property of allowing two kinds of really existing phenomena, those that exist independently of anyone’s thoughts or perceptions and those that exist independently of our (in the narrow sense), but not independently of others’ thoughts or perceptions.⁶ Human languages, being very fundamental forms of institutions, are of the second kind.

Fortunately, in the context of the present volume the issue of realism and its different opposing views does not really constitute a problem. Most of the authors seem to be realists about the world, although there may be considerable disagreement with respect to the degree to which different views on this world (alias ontologies) can diverge.⁷

1.2. Further dimensions of variation

Our short review of the depth dimension of the conceptual space that embeds different notions of ontology has brought to the fore the importance of the agents who have to do with ontologies (the plural makes it clear that the meta-level concept is meant). Here at least two roles have to be distinguished: The author of an ontology and its user. It is a trivial fact that they need not coincide, but an especially compelling illustration of this fact is provided by Rolf Pfeifer’s (2000) *Didabots*, simple robots for didactic purposes, who didn’t author but use an ontology which consists of only three situation categories: (a) no obstacle, (b) obstacle to the left, (c) obstacle to the right. *Didabots* have wheels and sensors and an algorithm that, based on their ontology, lets them avoid obstacles.⁸

Whereas the user of an ontology can always be identified, its author may well be unknown or even inexistent. Take *Caenorhabditis elegans*, the nematode or roundworm that became famous among other things for its nervous system which consists of only 302 neurons. One could say that the ontology of *C. elegans* comprises at least the following categories and their complements: (a) increasing concentration of an attractant, (b) decreasing concentration of a repellent, (c) increasing closeness to the preferred temperature. This makes sense insofar as the behavior of this nematode is geared towards situations of category (a), (b) and (c), and not their complements; the former two forms of behavior are called chemotaxis, the latter thermotaxis. So if one is ready to speak of the world view or ontology of a robot or a roundworm, the identity of the user is clear, but in the latter case the identity of the author is problematic. It therefore seems reasonable to assume that oftentimes ontologies have simply evolved, without any specific author available.

Alongside with the roles of the author and the user, the role of the object or domain of an ontology (in the right-hand column of our conceptual space) is of prime importance. It has already been shortly addressed in terms of the horizontal dimension of variation of our conceptual space, so a short reminder will suffice here. Ontologies vary with respect to their ‘aboutness’, i.e., what they are ontologies of. This can be, with decreasing generality, (i) all possible worlds, (ii) one world only, especially the one we live in (or seem to live in; cf. the depth dimension), (iii) subdomains of this world (or others) of increasing degrees of specificity. Is there an upper bound for the specificity in (iii)? In other words: What is the minimum degree of generality that is required for an ontology? Does it make sense to speak of the ontology of this nasty fly that keeps circling your head as you are reading this?

We submit that the answer should be negative. There is something like the ontology of *C. elegans*, or at least it is in the making,⁹ but this does not mean that this ontology provides a systematic account in all relevant aspects of a single exemplar of this species, but of all exemplars that come from the same kind of genome. Similarly, it certainly makes sense to develop an ontology not only of aircraft, but also one of aircraft accidents. But it does not make sense to create an ontology of the Airbus A 340 crash at Toronto on August 2nd, 2005,¹⁰ at least not without another significant extension of the concept. So far, even the most specific domain ontologies like the one of the famous roundworm have a generic object. They are intrinsically intensional insofar as they entail predictions about new entities that instantiate the generic entity. Insofar, ontological knowledge about a domain is definitional knowledge about it, not episodic knowledge about its states and fates.

The last and certainly not least aspect of ontologies that has to be addressed in this overview is the evaluation aspect. If there are two competing ontologies of the same domain, is it possible that one is true and the other one false? Or are there other criteria for the evaluation of ontologies? If, as we have assumed, ontologies in the simple meta-sense are conceptualizations, then they themselves cannot be true or false, they can only be more or less adequate and more or less useful. And of course they are completely inadequate if they entail false statements. If for instance someone conceptualizes human languages as being either red or green, this entails that English is either red or green, which is false, because it involves a category mistake. Category mistakes, especially less blunt ones, are a serious source for inadequacies in ontologies.

But even if an ontology does not entail false statements it can still be inadequate or less useful for various reasons. Usefulness is a relational concept which requires a purpose. Of two competing ontologies one can be more useful than the other for one purpose and less useful for another. Of two classifications of aircraft, e.g., one according to the kind of propulsion and the other one according to the status of the owner, the first one will outrank the latter in helpfulness when the purpose is spare parts, and the opposite will be the case when the purpose is legal matters of air traffic.

The usefulness of adequate conceptualizations and hence ontologies is a key issue in all scientific disciplines,¹¹ but it is especially important in disciplines that involve evolutionary accounts, such as biology, and even more so in fields where historical transfer plays a role, such as linguistics. Ontologies for linguistics are the topic of the contributions by Farrar and Zaefferer to this volume. Other aspects of the assessment of ontologies like provenance and credibility are discussed in Section 2.4.2.4. below.

Here is a summary of the findings of this section. In its most general reading, the article-less term *Ontology*, which lacks a proper plural, has turned out to refer to a rather controversial and indeed puzzling subfield of philosophy and more precisely of metaphysics. Thomas Hofweber, a philosopher of language, metaphysics, and mathematics, speaks of a “puzzle about ontology”, which he identifies as “the puzzle that there seem to be two contrary but equally good answers to the question (Q) How hard is it to answer ontological questions?”, namely “Answer I: Very hard” and “Answer II: Trivial.” (Hofweber 2005: 259). Still, the definition of the philosophical discipline as being concerned with “what entities make up reality” (Hofweber 2005: 256) seems to be relatively uncontroversial.

This is not the case with the notions the count noun *ontology*, with a proper and frequently used plural, is used to encode. Ontologies in this sense, specific answers to the question of what entities there are, may – but need not – be the outcome of philosophical endeavors, they may come from other disciplines as well or they may be no human artifacts at all, as in the case of the ontology of a macaque monkey brain (Metzinger and Gallese this vol.). In that third case the hardness question quoted above does not even arise (after all, it is the author of an ontology who has to answer ontological questions) and in the second case there is a clear tendency towards Answer II (authoring an aircraft spare parts ontology will rarely benefit from an ambition to answer deep questions). With respect to ontologies from non-philosophical disciplines, the controversy about the definition is mostly restricted to computer science (cf. Sections 2.3. and 2.4. below). The reason has to do with the prevailing conceptualization (or domain ontology) of language within that field. In principle, answering the question of what there is in a given domain may take any of various forms the outcome of a process of stocktaking may take: term lists, thesauri, glossaries or what have you. According to the traditional view, the linguistic items in these data structures are but strings of bytes and not full-fledged linguistic signs with form, structure, and content (cf. Farrar this vol.). So computer scientists tend to emphasize that a real ontology has to be much more than a mere inventory of items, i.e. lists of strings (cf. Section 2.3. below), and the controversy is mostly about what additional ingredients are required for an enriched inventory to count as an ontology.

The fact that philosophers never have thought about this kind of intricacies shows only that they always have taken the inferential potential of concepts for granted, whereas in computer science it takes a whole machinery to get it going.

We will conclude this section with Figure 3 that illustrates some of the dimensions of variation we have presented above. In the rows from top to bottom it shows very general, more special and very special variants of Ontology and ontologies, in the columns from left to right (i) the corresponding fields, (ii) examples for specific cases of accounting for what there is in a domain, and (iii) the purview or intended domain of such a specific account, and in the depth from front to back realist-objectivist and more idealist-subjectivist approaches. Not shown in the figure are the following aspects of ontologies we have also addressed briefly above: origin (author or evolutionary process), user, purpose and quality.

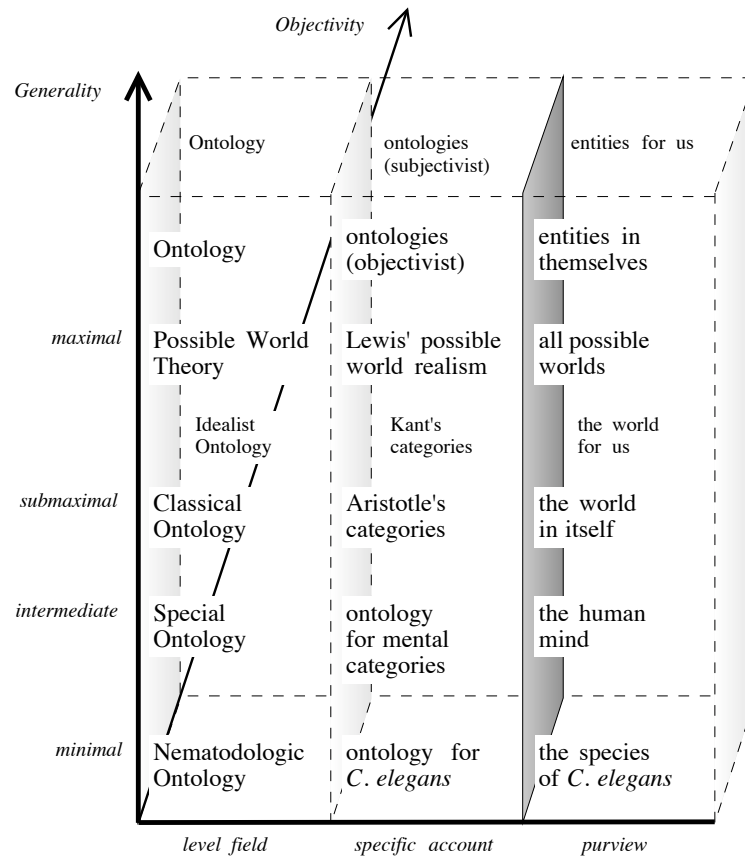


Figure 3. Example of variation dimensions and aspects of ontologies

2. Notions of ontology in different disciplines

2.1. Ontology and ontologies in other fields

Having carved up the conceptual space underlying the term 'ontology', we will now turn to a discussion of notions of ontology in different disciplines. This includes briefly touching on disciplines that are not directly relevant with respect to linguistics and thus not in the focus of this volume. Following this, we will concentrate on the disciplines and areas that are connected to the ontolinguistics enterprise, specifically on linguistics, computer science, and artificial intelligence.

Ontological questions are metaphysical questions by definition, and etymologically speaking the latter come after the physical questions, so they can be turned back on physical issues. Contemporary theoretical physics is full of ontological questions of the hard kind like the particle-wave duality, quantum ontology or the ontological implications of string theory, so Ontology as a discipline has interesting issues to deal with that come from physics. But physicists do not seem to make much use of ontologies of their domain.

The situation is different with the other hard sciences. One of the most successful ontologies of modern times seems to be the periodic table of elements in chemistry. This can be regarded as a paradigm case of an ontology not only because it is highly systematic and useful, but also for its predictiveness: At least two elements, those with atomic numbers 117 and 118, are claimed to exist, but have not yet been successfully attested. Another paradigmatic feature of this ontology is that it is about natural kinds.

The same has been assumed for a long time for the species of biology. But it turned out that although there is no doubt that there are different individual species (species *taxa* as biologists call them) like *Homo sapiens*, our species, or *Canis familiaris*, that of the domestic dog, it is doubtful, if taken together, that the species form a natural kind or category because there is no criterion that unites them to the exclusion of other taxa (interbreeding competes with common ecological niche and phylogenetic unity, cf. Ereshevsky 2002). Nevertheless, one of the domains that are currently characterized by an incredible boom in ontologies is that of the life sciences. Bioinformatics and related areas are teeming with web services like Open Biomedical Ontologies (“an umbrella web address for well-structured controlled vocabularies for shared use across different biological and medical domains,” <http://obo.sourceforge.net/>) or Gene Ontology (“a controlled vocabulary to describe gene and gene product attributes in any organism,” <http://www.geneontology.org/>).

Since the social sciences are only starting to discuss ontological issues (cf. our discussion of Searle’s contribution), it does not come as a surprise that social ontologies are not considered to be very mature so far, with a notable exception:

Legal systems are perhaps the most well-developed ontologies in the social world. Most laws are categorizations of objects at some level and most legal disputes turn on distinctions among categories. Because legal systems often comprise the most well-developed ontologies of the social world, they are a good reference for philosophers and social scientists seeking to study social objects. (Koepsell 1999: 219)

To conclude the general overview presented in this section, the increasing use of ontologies in business applications, e.g., workflow ontologies, should not remain unmentioned.

2.2. Notions of ontology in linguistics

Compared to other disciplines, contemporary linguistics assigns the term ontology a rather peripheral role in its domain. So far, only two subfields seem to make systematic use of one or the other member of the family of concepts coded by this term and a third one is just beginning to do so.

The younger one is that branch of computational linguistics that systematically takes advantage of ontologies in the AI sense of the term. One of the most recent and most extensive outcomes of this field is the book *Ontological Semantics* (Nirenburg and Raskin 2004). Its ontology lists the definitions of concepts for describing the meanings of lexical items of natural languages, but also for the specification of the meanings of the text-meaning representations that serve among others the function of an interlingua for machine translation. Other intended applications are information extraction, question answering, human-computer dialog systems, and text summarization.

The older linguistic domain that uses the term ontology is model-theoretic formal semantics, which in one of its simplest guises uses the ontology of first order logic: individuals, sets of tuples of individuals and truth values. But progress in the semantic analysis of natural language made it soon obvious that a less parsimonious ontology is required for this purpose and so the number of ontological categories started to grow. To point out just a few milestones: Montague (1973) added possible worlds and moments of time to the basic ontology and projected from there, using a recursively defined set of types, an infinite ontology of possible denotations; Link (1983) devised an integrated ontology for individuals and substances by providing both with a semilattice structure; Davidson's (1967) proposal to grant events the status of a basic ontological category has been welcomed and fruitfully employed in linguistics, and Barwise and Perry's idea (1983) of enriching the ontology by admitting situations as 'first-class citizens' had a similar impact. In short, many contributed to the task of freeing linguistic ontology from the constraints of philosophical ontology. In the abovementioned paper, Godehard Link advocated the view that "reductionist ontological considerations" are "quite alien to the purpose of logically analyzing the inference structures

of natural language” and went on to state the maxim: “Our guide in ontological matters has to be language itself, it seems to me.” (Link 1983: 303–304)

In the same spirit Emmon Bach coined the term ‘Natural Language Metaphysics’ (cf. Bach 1986; “. . . what I am doing here is not metaphysics *per se* but *natural language metaphysics*,” Bach 1989: 98). He is also to be credited with the most succinct characterization of the difference between the two:¹² Whereas the philosopher is interested in answering question (1), the job of the linguist is to find convincing answers to question (2):

- (1) *What kinds of things are there?*
- (2) *What kinds of things do people talk as if there are?*

We have called above the field that revolves around (1) Ontology, consequently we will use the term Language Ontology (with two capital initials) for the endeavors around (2), if language in general is concerned, and language Ontology, e.g., Korean language Ontology, when the focus is on an individual language.

It should be clear by now that the subject-matter of Language Ontology does not concern only model-theoretic semantics (it simply becomes visible there most clearly), but should interest every linguist who subscribes to the view that linguistic signs associate perceivable forms with conceptual contents, because these conceptual contents are never isolated in human language users, but integrated into the way they conceptualize their world, their individual ontology. Individual ontologies contain one or more language ontologies, but also something else which is often called commonsense ontology.

We submit that Commonsense Ontology, the study of commonsense ontologies, should be defined as being about answering the question (3):

- (3) *What kinds of things do agents behave as if there are?*

There are two reasons for distinguishing Language Ontology from Commonsense Ontology (and the same holds for their lower-case counterparts). The first is that it makes sense to ascribe commonsense ontologies also to subjects that do not have language (like robots, macaque brains, and roundworms), and the second is that the question of the relation between the two is too interesting to be begged by blunt stipulation of their identity. The pertinent keyword is linguistic relativity and the challenge consists in factoring the ontogenesis of individual ontologies into (a) the conceptual default settings babies are born with, (b) the culturally induced development the conceptual

system is subject to, and (c) the effects of the individual language on this development. This points to another discipline that uses the term ontology, not entirely within the confines of linguistics proper, but overlapping with it: developmental cognitive science.

A look into the child development literature shows lively research activities in this field. Imai and Gentner, e.g., tested whether the distinction between object names and substance names is based on a pre-linguistic ontological distinction or is driven by language (Japanese and English). Their results lead them to the following speculation:

Children begin learning word meanings building on their pre-linguistic ontological knowledge about individuation. Language learning leads children to pay attention to those aspects of the world that are habitually used in their own language, and this influence begins very early. Finally, children's sensitivity to linguistically-relevant aspects of the world may come to extend beyond the context of language use. (Imai and Gentner 1997: 196–197)

There is no doubt that language ontologies and commonsense ontologies are closely related since every language ontology is “a conceptualization or categorization of what normal everyday human language can talk about” (Zaefferer 2002: 33–34) and this is largely determined by the requirements of everyday life. In other words, both primarily contain concepts of entities encountered in everyday life (for an overview of what that could and should comprise, cf. Zaefferer this vol.) and their relations. These are the concepts for the expression of which natural languages tend to readily provide codings, be they simple or complex. Examples include concepts such as CAR or RUN, the most compact codings of which in English are the noun *car* and the verb *run* (more complex codings like *motorized vehicle with wheels* or *go faster than a walk* are reserved for special purposes), and relations such as the conceptual subordination of CAR under VEHICLE and of RUN under MOVE. Systems for the representation of word semantics such as WordNet (Fellbaum 1998, cf. also Fellbaum this vol.) are based on sense relations and thus reflect the underlying language ontology, since sense relations are relations between words (in a reading) based on ontological relations between the concepts that constitute the meanings of these words (in that reading).

Lexical semantics is not the only example of linguistic research that needs to take the corresponding language ontology into account. Another case in point is work on classifiers – be it in the context of classifier systems in different spoken languages (cf. again Imai and Gentner 1997 and Hellwig this

vol.) or of classifier predicates in signed languages (Talmy this vol.). Here, the underlying conceptual categorization of entities is responsible for the use of different classifier morphemes or predicates.

A further example is the almost trivial observation that different word classes tend to reflect different conceptualizations in that verbs usually code eventities whereas non-derived nouns most often code ‘thing-like’ entities (called ‘inventities’ in Zaefferer this vol., and ‘ineventities’ in Schalley 2004). Based on this, adjectives code mostly characteristics or attributes of the latter, whereas adverbs do the same for the former.

A fourth kind of linguistic studies where language ontology plays a role is anaphora resolution. In a sentence like *When you try to catch a lizard, the reptile may drop the tail and escape* both definite noun phrases are anaphorically related to the indefinite *a lizard*, but the relationship is mediated by ontological relations of different kinds (for the definitions compare Schalley and Zaefferer this vol.): Since LIZARD is c-subordinated to REPTILE (every lizard is a reptile), *the reptile* may have its antecedent in *a lizard*, and since LIZARD is m-i-superordinated to TAIL (every complete lizard has a tail as integral part), *the tail* may be interpreted as including a possessor slot which again has its antecedent in *a lizard*. Given the ontolinguistic framework the former case could be called conceptual subanaphor and the latter meronomic superanaphor.

Our final example showing the relevance of ontological knowledge for the proper use of language is the grammar of coordination. A precondition for the coordination of phrases as well as sentences is that the conjuncts are parallel with respect to syntax, semantics, and prosody (Lang 1984), where semantic parallelism is defined by two constraints: (a) the concepts coded by the coordinated elements have to be semantically independent, i.e., neither of them is c-subordinated to the other, and (b) there has to be a non-trivial subordinator, a third concept that is c-superordinated to both. So *my dog and my animal* and *a walk and an integer* are both semantically bad noun phrases, the first for violating (a), since DOG is c-subordinated to ANIMAL, and the second for violating (b), since the strongest common c-superordinate of WALK and INTEGER is probably ENTITY and so it could not be more trivial.

As mentioned at the beginning of this section there is a third subfield of linguistics that makes use of the term ontology or rather is just beginning to do so. It is the field of foundations of linguistic theory together with the field that deals with linguistic terminology. Whereas the former is concerned with questions various other disciplines like philosophy of language and philoso-

phy of science are also interested in, the latter has among other things to keep pace with all the terminological innovations that keep growing in the different schools of linguistics around the globe. And still they are united by a common interest in what will be called ‘Ontology for linguistics’ (or ‘ontologies for linguistics’) here in order to distinguish it from Language Ontology and its kin. Continuing our strategy we will characterize also this field by its leading question:

(4) *What kinds of things linguists talk about are there?*

Strictly speaking, since the study of linguistic terminology is about linguistic metalanguage, a variation of leading question (2) above, characterizing Language Ontology, would seem to be more adequate for this field, namely (5):

(5) *What kinds of things do linguists talk as if there are?*

However, given the scientific ambition of linguistics, a separate investigation of (5) without consideration of (4) will not be satisfactory. So the notion ‘ontology for linguistics’ refers to those conceptualizations of the domain of language and languages that are used to ‘talk linguistics’, to express and describe linguistic phenomena with the help of the corresponding concepts and the relations between them. The linguistic codings of these concepts are often, but by no means exclusively, technical terms of linguistics. Examples for such concepts include WORD CLASS, SPEECH ACT or EVENTITY (coded, e.g., in German linguistic terminology by the nouns *Wortart*, *Sprechakt* and *Eventität*, respectively), and CONCEPT (coded by the English word *concept* as used in this chapter), but also relations such as the conceptual incompatibility between ARTICLE and VERB (i.e., it is not conceivable that some linguistic entity is both an article and a verb).

Interestingly, a certain ontology of linguistics is also part of natural language ontology, since the codings of corresponding concepts such as SAY, WORD or QUESTION are presumably part of every natural language (cf. also Goddard this vol.; Zaefferer this vol.). Some of these concepts have a special status in any language ontology as they are instantiated by linguistic signs, which in turn have instantiations that refer to a concept. The concept WORD, e.g., is instantiated among others by the English word *word*, whose tokens have the potential to activate in hearers mental representations of the concept WORD.¹³ Since the same holds for many concepts coded by technical terms of linguistics, it is obvious that any ontology for linguistics has to include a

meta-language ontology, that is a language ontology for the language that is used to describe linguistic phenomena (cf. Farrar this vol. and Zaefferer this vol.).

There are at least two reasons why both explicitly spelled-out language ontologies and well-defined ontologies for linguistics are urgent desiderata in current linguistics. First, it is a truism that precise descriptions of linguistic phenomena without precisely defined technical terms are impossible. And second, only with the help of these tools can linguists reliably compare and compile different descriptions within a language and across languages. However, there are still many areas in linguistics that are characterized by confusion and disagreement on the terminology used – to pick just one example, the area of information structuring with all kinds of uses for words such as *topic* or *given* or *background* – and therefore it is often far from clear if different authors and schools presuppose different ontologies for linguistics or if only the labels vary.

Given these circumstances it is to be highly welcomed that projects like GOLD ('General Ontology for Linguistic Description', cf. Farrar this vol.) or DOLPhen ('Domain Ontology for Linguistic Phenomena', cf. Zaefferer this vol.) are on their way. Quite a few terminological problems that arise in linguistics are due to a lack of awareness of ontological differences. Consider sense relations for instance. Sense relations structure the lexicon in that they reflect conceptual relations that hold between the readings of lexemes. In talk about sense relations the distinction between conceptual relations and relations between the corresponding linguistic signs is often blurred or not drawn at all. Whereas the relation between CAR and CHASSIS is a conceptual one – a meronymic relation – the corresponding relation between any linguistic codings of those concepts, e.g. between the English nouns *car* and *chassis*, is a semantic relation – a meronymic one. Another example would be the hyponymy relation between the English nouns *car* and *vehicle*, which holds because of the conceptual subordination of CAR under VEHICLE. Yet, the role of the corresponding language ontology is typically left implicit, in that criteria for the sense relations are formulated in terms of linguistic characteristics or 'meaning' (cf. Cruse 1986, for instance, or Schalley 2004: 27–29), but not in ontological terms.

If meanings are just concepts that happen to be coded by a given meaning-bearing entity of a given language, for instance a word, then of course meaning relations are just ontological relations. But linguists are rarely aware of the fact that relations of this kind hold irrespectively of how the related con-

cepts are coded and irrespectively of whether they are coded (and therefore meanings) at all. It is as if people would speak of manned space capsule *A* being in love with manned space capsule *B* when what they mean is that the man aboard *A* is in love with the woman aboard *B*.

In summary, it appears that, in order to improve the analysis of linguistic facts, linguists would need to give yet more attention and weight to the study of underlying conceptualizations (including their interconceptual relations) both in the users of their object languages and in themselves as users of the linguistic metalanguage. This area seems to have the potential for considerable progress through explicit and systematic investigation of the language-ontology interface. The present volume aims to put corresponding current efforts into a broader context and to instigate a more systematic approach to ontologies in general, and to language ontologies as well as ontologies for linguistics in particular, by promoting an ontology-driven approach to linguistics and thus by arguing for and exemplifying what we are calling ontolinguistics.

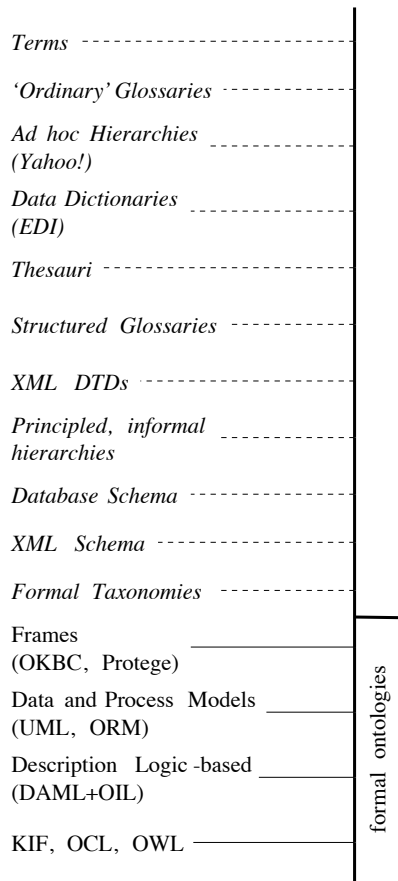
Given that the construction and maintenance of ontologies by hand becomes quickly cumbersome with increasing size, it seems reasonable to consider using corresponding tools from computer science, and therefore the next section presents an outline of the state of the art in the relevant subsections of this thriving discipline.

2.3. Ontologies in computer science: A survey

As indicated above, the term ontology is widely used by the computer science community and there it refers broadly speaking to the construction of information models. In computer science an information model is an abstract formal representation of entities that includes their properties and the relations that hold between them. By contrast with data models, information models represent their entities without any specification of implementation issues or protocols for data transportation. Among computer scientists the word ontology has received such a broad use that it has been employed to refer to any information model. It is necessary for the purpose of the following to constrain that usage. More specifically, we will understand by ontology in the computer science sense a specification in a formal language of terms and definitions describing things that make up the world.

A key component of this definition of ontology is the phrase ‘a formal language’: Different degrees of formality are exhibited by different informa-

tion modeling languages. Figure 4 presents a set of such modeling languages along a continuum (excerpted from Ray 2004). We will also briefly describe them in order to delineate what counts as ontology in computer science.



Terms refers to a controlled and usually domain specific vocabulary. ‘*Ordinary*’ *Glossaries* are terms with natural language definitions, as one finds in many textbooks, *Ad hoc Hierarchies* such as Yahoo! are sets of terms with a relationship between terms, but where no formal semantics for that relationship is defined. *Data Dictionaries* are more formal models of information, often of relational databases, where each term and relation has an associated natural language definition (EDI stands for standardized Electronic Data Interchange). Since Roget’s pioneering work a *Thesaurus* is a word list that is not ordered alphabetically but according to conceptual relations. *Structured Glossaries* may include relationships among the terms in the glossary. *XML DTDs* are Document Type Definitions in eXtensible Markup Language (Yergeau et al. 2004), used for communication among software systems. XML

Figure 4. Information modeling languages information structures, but is a language for defining syntax that has no associated constraints on semantics. *Principled Informal Hierarchies* are those which do not have a formal, logical model of relations between terms, but at least have an informal, common-sense explanation of relationships. *DB Schemas* are Data Base structures that have more formal definitions of the meaning of terms and relations, usually by employing statements in a database constraint language. *XML Schema* (with a capital S) is a further

development of the XML DTD and is now the way to specify XML-based communication that is recommended by the World Wide Web Consortium. *Formal Taxonomies* are those which have a formal, logical semantics for the relations among terms.

Frames include a range of standard AI languages that have terms, relations, and inheritance of properties. Examples are the Open Knowledge Base Connectivity (OKBC) protocol (Chaudhri et al. 1998) and the ontology editor and knowledge acquisition system Protégé. *Data and Process Models* couple taxonomies and defined relationships with a semantics for representing process and action. UML, the Unified Modeling Language for specifying the design of object-oriented systems (Object Management Group 1997–2006, cf. also Schalley this vol.) and Object-Role Modeling (ORM) exemplify this kind of information model. *Description Logic* (Baader et al. 2003) languages like the DARPA Agent Markup Language (DAML), the Ontology Integration Language (OIL), and their merger DAML+OIL combine the knowledge representation elements of a frame system with the ability to define rules; they are sublanguages of predicate logic. The expressiveness of rules is limited in order to ensure that inference on the rules is tractable. *KIF*, *OCL*, and *OWL* refer to three very expressive languages. Knowledge Interchange Format (KIF) (Genesereth 1991) is a first order logic, for which there have been several versions that differ in their details. Object Constraint Language (OCL) is part of UML (see above). The Ontology Web Language (OWL) (McGuinness and van Harmelen 2004) is a formal logical language, with similar expressiveness to KIF, that conforms to XML syntax. It should be noted that there are languages with far higher expressiveness including modal logic and various higher-order logics (Nadathur and Miller 1998) that exist further down along the continuum shown in Figure 4. They present significant challenges for practical inference however, and to date they have been used primarily for research in theorem proving rather than specification of ontologies.

It is only models that make use of the full features of the languages in the bottom of the diagram, from ‘Frames’ onward, that can be called ontologies in a way consistent with our definition. In order to distinguish ontologies in computer science (and artificial intelligence) – which are based on a formal language – from ontologies in a more general sense, we will refer to the former as ‘formal ontology’ hereafter.

A formal ontology is distinct from the most common instance of a set of terms and definitions: the dictionary. A dictionary does not employ a formal language, but rather an informal one: a human natural language. A dictio-

nary is meant to be read and interpreted by humans. No machine is currently capable of understanding a dictionary in any realistic sense of the word ‘understanding’. Furthermore, a dictionary is descriptive. It provides definitions which are presumably appropriate at a point in time, often with annotations about the usage of words up to the time of publication. Language evolves through an organic process and all attempts to render language static, such as the efforts of the *Académie Française*, have failed and will continue to fail to a significant degree. In contrast, a formal ontology is prescriptive or normative. It states definitively what a given term means in a formal language. A term in an ontology is not a word but a concept, although the concept will normally be given a name which is a word or combination of words in order to support human understanding of the ontology. A true formal ontology however could have all its term names replaced with arbitrary codes and still have the same formal properties. The only issue would be how such an ontology relates its terms to linguistic items in order to make its results of processing intelligible and useful to humans.

One of the issues within the ontolinguistics enterprise is indeed this: how can relations between ontologies – as they are used in computer science – and linguistic expressions be established? Although any such relation will be imperfect, the degree of precision of relation and scope of coverage has been improving since greater bodies of formal ontologies, lexical resources, and corpora became available. The most prominent lexical database is WordNet (Fellbaum 1998), and there are efforts to create similar resources in languages other than English, often with relation to the English WordNet. Such resources focus on the smallest lexical units, which are usually words, although multi-word units are also present in small numbers (cf. Fellbaum this vol.). Collections of larger, phrasal units have been proposed (Pease and Fellbaum 2004) and the collection of lexical functions proposed by Mel’cuk (1998) – universal relations between lexical items including the standard sense relations – has been studied. There is significant potential for the ontology community to make use of work undertaken to catalog closed-class elements of language. Such elements may be considered to have a significant place in communication due to their presence as structural features in languages, as opposed to the elements of the open-class or lexical subsystem (cf. also Talmy this vol.). A more recent effort to relate a formal ontology to WordNet (Niles and Pease 2003) is also described in this volume (Pease this vol.).

However, lexical resources and formal ontologies are very different artifacts. For instance, over the past few years there have been many publica-

tions that describe ‘fixes’ to the WordNet taxonomy according to ontological principles (Gangemi et al. 2002b). Fundamentally, these are misguided, since language, as an organic system, does not conform to ontological principles. Once this distinction is recognized, however, there is great value in relating language to ontology for use in a broad range of applications and research endeavors. One innovative effort in this volume (Farrar this vol.) makes use of a formal ontology to describe structural linguistic information itself, an approach that has been touched on from a linguistic perspective in the previous section.

There is an important distinction in Formal Ontology (and also Ontology more generally, as has extensively been discussed in Section 1.) between the language in which an ontology is expressed and the content or semantics of the ontology itself. A much larger proportion of effort in the Ontology community in computer science has gone into the development of languages as well as tools and methods, compared to the level of effort that has gone into the creation of content.

One aspect connected to the creation of content is the semantic scope and degree of generality an ontology exhibits (cf. also the discussion in Section 1.). This naturally applies to formal ontologies as well – which are characterizable by whether they cover very general concepts, as in an upper ontology, or very specific topics, as in a domain specific ontology. Most extant formal ontologies pertain to fairly narrow topics or domains (Casati and Varzi 1995; Grüninger and Menzel 2003), although these can be of great interest and value. The authors are aware of only three formal ontologies that have attempted to define the broadest and most general notions, which collectively may be termed an upper ontology. These are the Suggested Upper Merged Ontology (Niles and Pease 2001), Cyc (Lenat 1995), and DOLCE (Gangemi et al. 2002a). We should note that the distinction between upper ontology and domain specific ontology, however, is a continuum without a clear dividing line. While the relation `temporallyBefore` is certainly an upper ontology concept and the class `Carburator` is certainly a domain specific one, there are many concepts in between such extremes that do not have such an obvious membership.

Another way of characterizing formal ontologies is the number of terms in the ontology. Terms may be classified into, amongst others,

- instances, like `KofiAnnan` and `Germany`;
- classes, like `Human`, and `Country`;

- relations, like `agentOf` and `mother`;
- function terms, like `GovernmentOf` and `AdditionFunction`.

However, a count of terms can only be a meaningful metric as to the size of an ontology, if the terms counted include significant associated definitions (through which the terms are sufficiently interrelated).

Yet another metric is the number of axioms, which are indispensable for inference and expressive power. An axiom is any formal statement. Such statements may be

- simple ground statements, like ‘Kofi Annan is Secretary General of the UN’, and ‘Germany is a country’;
- quantified statements, like ‘there exists some farmer who beats his donkey’;
- rules, like ‘Every good boy loves his mother’.

Note that here we have stated in informal English examples that would be expressed in a formal language in a formal ontology.

2.4. Ontologies and artificial intelligence

According to a prevalent definition by Luger and Stubblefield (1993), Artificial Intelligence (AI) is the branch of computer science that is concerned with the automation of intelligent behavior. Regardless whether this effort is based on the abstract concept of rationality or involves mimicking the human mind, any truly intelligent computer system surely requires the capability to acquire, process and use knowledge about the domain it is situated in or concerned with. Whereas until the early 80s, this capability was usually associated with the presence of some knowledge storage facility and logical reasoning (like in so-called *expert systems*), nowadays it is no longer consensus in AI research that information processing demands for an *explicit* collection of computational knowledge (a so-called *knowledge base*), possibly including some domain conceptualization in form of a formal ontology. Approaches in some contemporary AI-subfields such as *connectionism* or *situated intelligence* usually lack any symbolic knowledge representation, and certain popular AI methods such as *Q-learning* get along without any (explicit) world model. Nevertheless, the presence of a knowledge base (whether based on logic or another knowledge representation format such as

Bayesian networks) can strongly increase the flexibility and adaptability of an intelligent system by virtue of the separation of knowledge collection and knowledge-based reasoning on the one hand, and planning, searching and acting capabilities on the other, as pointed out in detail in Section 2.4.1. In addition, the ability to represent and process information symbolically can be a prerequisite for the exchange of information with others (see below). Consequently, the majority of AI-frameworks still comprise some kind of more or less powerful knowledge base, and an increasing number also provides some facility specifically dedicated to concept knowledge. The latter is the case for basically the same reasons why ontologies are used in other disciplines and in “ordinary” computer science too (e.g., to be able to reuse rather general, abstract and persistent domain theories in different tasks of knowledge processing). But in contrast to other areas of computer science, in AI the foci are on the computational reasoning about/using (ontological) knowledge, on the (computationally) intelligent acquisition and revision of new knowledge, and on the computational use of (ontological) knowledge for decision making. In addition, ontologies in AI are very useful and often inevitable means for the communication and collaboration of intelligent systems, including the interaction of humans with machines and the machine-supported interaction of humans. In such settings a common informational ground needs to be found in order to facilitate understanding and cooperation (and even conflicts), and consented ontologies partially provide this common ground in terms of conceptualizations of the respective common domain. This states a reason for the eminent importance of ontologies in Distributed AI, a fast-growing subfield of AI concerned with the interaction of intelligent systems, as described in detail in Section 2.4.2.

2.4.1. Knowledge bases

The notion of a knowledge base, or a collection of facts and rules, along with inference procedures to make use of those rules, has a long history in the field of AI (Pease, Liuzzi, and Gunning 2001). The point of this research area has been to decouple declaratively specified knowledge from procedural code, allowing a software system to behave more intelligently, and less mechanically, by dynamically combining small chunks of knowledge to reach an answer. Whereas a conventional software system would have specified a series of operations to be performed in a certain order, a knowledge base sys-

tem has a generic inference process that can opportunistically apply a range of declaratively specified knowledge in order to reach different answers to different queries.

Knowledge base systems are also called expert systems in part because this work was primarily undertaken to provide a software-based expert in some field. One of the earliest expert systems was Mycin (Buchanan and Shortliffe 1984), which was designed to diagnose infectious blood diseases. It achieved a level of competence that was better than most human experts. Notably, despite its competence, it was not put into commercial use because of social and human factor issues. This, in part, spawned a whole new field of research to address these ‘soft’ considerations in the successful application of software systems. Mycin did not originally have a completely clear separation of knowledge from inference procedures. The Emycin project (van Melle 1982) was an attempt to make that separation clear by creating a more general expert system shell that could be used on a more widely varied set of knowledge. Many companies created and sold expert system shells, some derived directly from the Emycin work.

As expert system projects proliferated in the 1980s, it became clear that although inference processes could be reused, knowledge often had to be recreated for each new application. The problem was that simplifying assumptions were often built into the knowledge. These assumptions were invariably appropriate for one domain, but not another. One way of interpreting this larger issue of assumptions was called the *frame problem* (McCarthy and Hayes 1969). The real world is large and complex. It is not practical to model every feature of the world, especially when a project has a focused goal such as diagnosing blood diseases. This tension between the need to focus knowledge creation on the task at hand, and to make that knowledge as reusable as possible, has spawned the Ontology sub-field of AI.

One head-on approach to this problem has been to attempt to encode all the common sense knowledge of the world (Lenat 1995). More modest efforts have been to create knowledge that is at a level of generality and reusability that is simply greater than most applications. Creating such knowledge however has proven difficult, and there are only three major attempts to create formal upper level ontologies (cf. above). Most research in this area has focused on tools such as *Protégé* (Eriksson et al. 1999), languages such as *KIF* (Genesereth 1991) and *OWL* (McGuinness and van Harmelen 2004), and processes as described in Guarino and Welty (2002), rather than on the knowledge itself.

2.4.2. *Ontologies in distributed AI: Issues and selected approaches*

The last few years have seen a tremendous rise of interest in ontologies for the use in *distributed settings* with multiple, interacting participants – especially so-called *open environments* like the Semantic Web, open multiagent systems and peer-to-peer systems.

Such environments can be characterized by the following properties, which might be more or less distinct depending on the concrete application domain:

- Heterogeneous set of autonomous participants, with only few restrictions for participation.
The participants operate basically self-interested towards their individual and often hidden goals. Neither the concrete set of participants nor their capabilities, beliefs, and intentions are known beforehand.
- The knowledge domain is highly dynamic and heterogeneous.
- Initial and possibly persistent nonexistence of a commonly agreed, single ‘truth’ (and thus of ‘knowledge’ in an objectivist sense), absence of a central instance for the enactment of behavioral and informational norms.

Applications in such environments require a shared domain semantics in order to support a mutual understanding among the distributed participants. Computational ontologies constitute a popular response to this need. The use of ontologies in such settings focuses thus mainly on the enabling of knowledge *communication*, *sharing*, and *reuse* by means of the generation and provision of a conceptual common ground among the interacting parties. In this regard, we distinguish two probably overlapping classes of participants, namely *ontology sources* and *users*, both human as well as artificial agents, but also ‘passive’ ontology sources like web documents.

As stated earlier, ontologies are usually defined as formal representations of domain conceptualizations, focusing on *consented* and *stable* concepts. In the following, we will describe issues which arise in open environments from applying this traditional understanding of ontologies, and present selected research efforts in response to these issues. For lack of space, we cannot give an exhaustive overview of the research field. Rather we would like to provide a short, concise description of key properties of a relatively small group of formal and technical frameworks, selected from a large list (with often equally useful other approaches), in order to provide a starting point for further reading.

2.4.2.1. *Issues*

While the compilation, integration, and sharing of information among heterogeneous information sources and users becomes more and more important, a large number of contemporary frameworks and tools for the modeling and usage of ontologies still proceed on the assumption of static, fully consented and authoritative ontologies. This reflects a sort of dilemma: On the one hand, ontologies should ease the reuse of knowledge and the sharing of knowledge among distributed parties, and thus should be stable and agreed, on the other hand, ontologies (being a special kind of knowledge) are themselves subject to difficulties known from the field of knowledge sharing, e.g. arising from controversial viewpoints.

The following problems are considered to be most prominent due to distributed settings with heterogeneous ontology sources. Further potential issues and examples can be found in Tamma (2002) and Staab and Studer (2003).

1. *Ontology sources operating with mutually incompatible representation languages.* This issue is likely to become less severe in the near future, since the standardization of representational aspects progresses rapidly, mainly driven by the Semantic Web effort. But there still is no consensus regarding the adequate degree of expressiveness required for a common formal ontology language (taking into account aspects like logical decidability). This is one reason why the W3C's *description logic* based ontology language *OWL* (cf. Section 2.3.) comes in three variations – *OWL Lite*, *OWL-DL* (the most often used variant of OWL which is equivalent to a certain prevalent description logic), and *OWL Full* – differing in expressiveness, not to mention languages such as *SWRL* that extends OWL with the ability to represent rules. OWL-DL can be considered as the current quasi-standard for the representation of web ontologies, since it is considerably more expressive than OWL Lite while corresponding (in contrast to OWL Full) to a decidable variant of description logic.
2. *Homonymy and synonymy.* E.g., i) the same name is used for different concepts (because of context-dependency of the name, for instance, as with the word *wood* which denotes both a collection of trees and their material), or ii) different names are used for the same concept, like *car* and *automobile* (Tamma 2002).
3. *Incompatible concept coverage, scopes and modeling granularities:* E.g., i) multiple concept definitions appear to describe the same concepts, but overlap only partially in fact (for instance, in one ontology, CHAMPAGNE

might be a sub-class of WINE, in another it is not), or ii) concepts are modeled in a fine-grained way (i.e., with many sub-classes and/or attributes) in one ontology, but only coarsely in another (for instance, RED WINE and WHITE WINE as the only sub-class of WINE vs. CHARDONNAY, CHIANTI, BEAUJOLAIS, ... [Tamma 2002]).

4. *Incompatible representation paradigms and top-level concepts.* E.g., ontology 1 might use the elementary concept EVENT as a top node, whereas ontology 2 subsumes everything under MATTER.
5. *Semantic inconsistencies due to stable goal or belief conflicts of the participants.* E.g., the first ontology source treats RELIGION as a direct aspect of CULTURE, whereas the second contains an additional concept SUPERSTITION which is used to classify RELIGION. While this is an extreme example, one can easily find further examples from various controversial fields like politics.
6. *Communication problems, preventing agreement on and coordination of ontology sources; unreliable ontology sources.*

Whereas some of these issues can be resolved at least in principle (e.g., incompatible ontologies could be made compatible just by unifying the names of certain concepts, as with issue 2 ii) above), some other issues might be technically very difficult to resolve (e.g., issues 1 and 3), or the dissolution is impossible on the level of ontology processing at least without disproportional measures that would lead to severe restrictions on the software applicability. For example, semantically inconsistent definitions of the same concept might have their origin in divergent world views of the ontology sources (issue 5); an alignment of these world views, establishing an agreement, would – if practicable at all – lead to a loss in source autonomy and therefore decrease the flexibility and robustness of the application.

In addition to the possibility that some of the above issues cannot be resolved, and the possible undesired loss of ontology source autonomy when attempting to get rid of them, there are further considerations to be taken into account when it comes to the integration of ontologies in open environments which have been largely neglected in traditional approaches to ontology integration and sharing.

First, stable semantic conflicts are not just something one should get rid of by any means. Instead, *conflict knowledge* (Müller and Dieng 2002) (i.e., meta-knowledge about conflicts) can provide valuable information about the attitudes, world views and goals of the respective knowledge sources. More generally, a set of distributed ontology sources and users forms a *social layer*

consisting of provenance information and information about the social relationships among ontology contributors and users. The explication and evaluation of this layer can provide the knowledge users with valuable meta-knowledge, and – if made explicit and visible – can be prerequisites for a subsequent resolution of conflicts regarding controversial knowledge. In this regard, it is important to bear in mind that subjective intentions and goals do not just exist for intelligent agents, but also indirectly for other kinds of ontology sources (like web documents), simply by virtue of their human creators.

Second, in the absence of normative meaning governance, and due to the inherently dynamic nature of knowledge in open environments, such mechanisms for ontology integration as filtering, use of trust relationships and most traditional ways of ontology merging can only provide preliminary decisions about the reasonable modeling of domains, because within a heterogeneous group of autonomous ontology sources and users, in the end each user can only decide for himself about the meaning, relevance and correctness of the given information, and these decisions might need to be revised in the course of time.

2.4.2.2. *Ontology integration*

To integrate data from multiple ontologies, there exist different possibilities, namely ontology *merging*, *mapping*, and *matching* (Gruber 1993; Tamma 2002; Staab and Studer 2003). Whereas merging describes the process of creating a new single coherent ontology that includes the information of all merged ontologies, mapping describes a process where the original ontologies remain separated, but are made consistent and coherent with each other, either by finding tuples of related concepts and/or by defining mappings to relate concepts within the source ontologies. ‘Matching’ in particular deals with the problem of finding equalities or at least similarities among several ontologies. The central problem for the process of mapping/matching is to identify and compare the meaning of the respective concepts. Subsequently, a merging process can then unite equivalent concept descriptions, and remove redundant ones. There exist several clues for this task, e.g., from the comparison of the concept names to the evaluation and comparison of the sub-concepts of a concept and their relations.

Even today such processes are still largely conducted by hand, which is a time-consuming process that often leads to mistakes. Especially the fast growing number of distributed ontologies in the Semantic Web will therefore

increase the need for (semi-)automated tools in order to support ontology integration. There already exist a number of tools – such as, for instance, *OntoMerge* (Dou, McDermott, and Qi 2002), *OntoMorph* (Chalupsky 2000), or *Observer* (Mena et al. 2000) – supporting automatic or semi-automatic ontology merging or mapping. Their main purpose is to (possibly interactively) guide the ontology designer through the merging/mapping process, and to identify inconsistencies or other problems, and to present suggestions for further proceeding.

Advanced ontology and knowledge modeling environments which support multiple users are, e.g., *OntoEdit* (<http://www.ontoknowledge.org/tools/ontoedit.shtml>), *Protégé* (Eriksson et al. 1999), and *Sigma* (Pease 2003). Analogously, tools exist for the multi-user creation of ontology instances by means of the manual annotation of documents and other data with ontology-based meta-data. Of course, such approaches are of limited value if it comes to the integration of information in large-scale environments like the web. Software frameworks have been designed in order to support both the expert development of ontologies and the virtual and/or transformational integration and dissemination of heterogeneous ontologies and instance knowledge. A pioneering approach in this regard has been *OntoBroker* (Decker et al. 1999). A more recent, ambitious example for such a framework is *KAON* (Bozsak et al. 2002), which integrates available resources and provides tools for the acquisition, engineering, management, and presentation of distributed ontologies and meta-data.

If agents are involved in integration frameworks, they are often ‘only’ part of the technical middleware (serving as matchmakers, for instance), rather than being intelligent ontology sources themselves. One of few exceptions in this regard is *InfoSleuth* (Nodine, Fowler, and Perry 1999), where information agents provide ontologies that are used as media for the integration of heterogeneous knowledge contributions. Beside other features, InfoSleuth allows the annotation of knowledge contributions with information about their provenances and the long-term monitoring of knowledge domains.

One of the most important application fields for distributed computational ontologies is *Organizational Knowledge Management*, which addresses the ontology-supported creation, representation, usage, and distribution of knowledge within complex organizations, such as (large) companies or within the government. Most organizational knowledge management systems still aim for the creation of monolithic, centralized, and homogeneous knowledge bases for the collection of corporate knowledge, according to a

single ontology-based organization schema, in order to enable communication and knowledge sharing across the organization. An example for an organizational, agent-based knowledge management framework which in contrast explicitly acknowledges the distributed and social nature of knowledge in large organizations is *FRODO* (van Elst et al. 2004). *FRODO* can be characterized as a large-scale meta-knowledge system with ontology-based organizational structures and support for workflow-based knowledge contexts. It makes use of social agents for the management of ontologies, workflow, and personal information assistance in order to relate individuals and organizational concerns.

2.4.2.3. *Ontology emergence and uncertainty*

A characteristic of open environments is that knowledge domains are dynamic and can often be modeled with some uncertainty. *Probabilistic ontologies* (e.g., Giugno and Lukasiewicz 2002) provide the possibility to describe concepts and instances with variable degrees of belief, denoting *uncertainty* of description logic's terminological axioms (as opposed to *vagueness* in fuzzy logic). E.g., in a probabilistic ontology, the modeler can assign the probability 0.3 to the claim that 'Tomatoes are fruits'. Probabilistic ontologies usually build upon description logic, which can be used to describe assertional knowledge (i.e., about concept instances) also. Besides this, there are many other approaches for the probabilistic enhancement of knowledge bases which could in principle be used to model uncertain ontologies also (e.g., *stochastic logic*).

The *Simple HTML Ontology Extensions* (SHOE) (Heflin and Hendler 2000) is a quite fundamental approach to *dynamic ontologies*, acknowledging that knowledge on the internet is not static but evolves in time, and that ontologies do not exist in monolithic isolation. SHOE provides a formal framework and an ontology-based knowledge representation language intended for information embedded within web pages (via semantic annotations). It supports ontology revision as a change in the components of an ontology (i.e., the addition or removal of categories and their relationships), and the versioning of subsequently revised ontologies. In this regard, formal techniques like those described before are supported to align and integrate multiple ontologies.

A wide research field within that of ontology emergence is that of *ontology learning* from large unstructured or semi-structured data sets and natural-

language documents. Concrete techniques trade under names like *concept mining* and *clustering*, and build upon well-explored approaches in the areas of data mining and natural language processing. Mostly, they are limited to the automatic generation of taxonomical ontologies or word lists, though. Since this topic is beyond the scope of this article, we refer the interested reader to Staab et al. (2002) and Misskof, Navigli, and Velardi (2002) for details.

Almost all approaches to ontology sharing require an agreement on the respective concepts. In case such an agreement is not given, a process of establishing a common ontological ground among the parties has to be executed. In case these parties are intelligent software agents, such a process can be performed automatically: Ontology negotiation (Bailin and Truszkowski 2003) enables intelligent agents to cooperate in performing a task, even if their domain knowledge is based on different ontologies. Ontology negotiation allows agents to discover ontological conflicts and to establish a common ground for further communications through incremental mutual requests and interpretations, clarifications, and explanations regarding concept meanings. For this purpose, practical approaches to ontology negotiation usually provide an ontology negotiation protocol and a software infrastructure in order to support the negotiation tasks. Within this protocol, certain speech acts according to the negotiation tasks can be performed by the agents, like ‘Request Clarification’ (of an unknown concept name) and ‘Confirmation of Interpretation’ (of a given concept definition). In the course of the negotiation process, ideally, the agents come to an agreed categorization that can subsequently be used in knowledge-based communication. Due to its high communicational overhead it is questionable if this approach can be applied in the large scale, but it is surely a very flexible way for ontology alignment in dyadic micro-scenarios. Ontology negotiation acknowledges that different agents might have different world views that need to be aligned communicatively in order to facilitate meaningful further communication. Although this is a technical process steered by a rather simple protocol, this approach is somewhat related to that of linguistic ontological mediation (Bateman this vol.), which contends that besides perception-related ‘world ontologies’ of commonsense concepts there also exist a ‘linguistic ontology’ domain resulting from the construction of ‘reality’ by means of language. It is also related to the concept of purely socially (i.e., communicatively) constructed open ontologies (cf. below).

2.4.2.4. *Ontology assessment*

Ahead of any acquisition of ontological knowledge from an external source comes the rating of this knowledge regarding criteria like its credibility. This is especially important in case a selection is required from a set of inconsistent ontologies which cannot be integrated, or in case ontologies are found in an open environment where their providers cannot be trusted per se.

In order to assess the usability of an ontology, ontology-inherent properties like the containment of a certain category and its topical appropriateness have to be examined (which is not dealt with here, since it is not specific to distributed settings), but also external characteristics. The certainly most important meta-knowledge of such kind is that of the *provenance* of an ontology (or a part of it).

The co-presence of syntactically and semantically heterogeneous and even inconsistent ontological knowledge by virtue of provenance annotations is supported for instance by the *Web Knowledge Base 2* (WebKB-2) (Martin and Eklund 2001). The WebKB-2 server permits web users to add ontological and assertional knowledge to a shared, central base, such that syntactical and semantic heterogeneity is advocated to permit the comparison and mutual completion of knowledge proposed by heterogeneous ontology sources and users. The WebKB-2 has been initialized with *WordNet* (Miller 1995; Fellbaum this vol.) and a top-level ontology in order to provide initial content and guidance for its users. But the WebKB-2 still needs to prove its usefulness for real-world application, and currently does not have significant facilities for the comparison of its ‘inhabiting’, possibly highly heterogeneous part-ontologies.

Provenance information of ontologies is also provided by *Swoogle* (<http://swoogle.umbc.edu/>), which is an internet search machine specialized on OWL and RDF encoded formal ontologies and meta-data (RDF, the *Resource Description Framework*, is the currently most widely used language for the representation of information about resources on the web). It uses the *OntoRank* algorithm to rate multiple ontologies containing certain keywords. *OntoRank* is basically an adoption of Google’s famous *PageRank* algorithm.

Starting from provenance information, it is often possible to assign a measure for the reliability of ontology sources. *Trust* is a very broad and common notion in Distributed AI, usually expressing whether some ‘positive’ behavior can be anticipated, and to which degree. As for ontologies, trust information can be used to calculate degrees of beliefs in ontological statements (or any

other kind of claims) from the degree of trust assigned to the sources of these statements (Richardson, Agrawal, and Domingos 2003). As an alternative to the assignment of degrees of trust to knowledge sources, trust information can also be assigned directly to categorial and assertional statements (Fischer and Nickles to appear), which might allow for a more fine grained, context-aware trust management.

Although the concept of trust is widely used in knowledge modeling, and has an intuitive appeal, it should be used with care. Very often, ‘trust’ in a certain information source is plainly identified with the belief in every statement from this source, which is surely a much too simple understanding. Another issue especially in large application scenarios like the Semantic Web is that an efficient trust-based rating of knowledge would require an existing trust infrastructure like *trust networks* (Golbeck, Parsia, and Hendler 2003).

In general, any kind of ranking of information in terms of quality and credibility comes to its limits in case there is not yet enough information about trust or recommendations to identify and filter out ‘inappropriate’ or ‘wrong’ contributions, or there does not even exist an abstract concept of global inappropriateness or correctness at all. An alternative approach contrasting the identification and removal of semantic heterogeneity within or among ontologies is to maintain inconsistencies while integrating them, and to *reify* them using meta-knowledge about their provenances, their degrees of agreement by various knowledge sources, users, and groups (by virtue of voting on knowledge), and their personal and social contexts. Such *open ontologies* (Froehner, Nickles, and Weiss 2004) account for the fact that in open environments even ontologies (traditionally assumed to be more ‘objective’ and stable than other kinds of knowledge) are subject to social acceptance or rejection, and thus need to accommodate possibly divergent preferences and multiple points of view (e.g., controversial opinions) (Nickles et al. 2005) and optional mechanisms for the leveled fusion of heterogeneous ontological and ‘ordinary’ knowledge. Therefore, the focus is here not on the emergence of an agreement on a conceptualization, but on the provision of meta-knowledge about the personal and social circumstances steering the generation, propagation and usage of ontological knowledge.

In summary, distributed settings pose various challenges for the acquisition, representation, and usage of ontologies, caused by factors like information source autonomy, information heterogeneity, and the possible absence of commonly agreed conceptual knowledge in such settings. Approaches to computational ontologies need to address potential problems arising from

these factors, like inconsistent and un-trustable information, in order to be useful in open environments like the Semantic Web. Some common research approaches to the described issues have been outlined in this section. Very recent developments in the research of ontologies for open environments investigate *meta-modeling* as a technique to allow context-sensitivity of concept specifications (Motik 2005) (in the tradition of higher-order logic and *context logic* [McCarthy 1987]), a logic for the representation of possibly disagreeing opinions and other public attitudes (Fischer and Nickles to appear), and reasoning with *inconsistent ontologies* (Huang, van Harmelen, and ten Teije 2005). Some of these techniques have been known in Artificial Intelligence for quite a long time, but their practical use in very large and complex environments like the Semantic Web requires major adaptations, e.g. to ensure logic decidability.

Notes

1. The full frontispiece reads OGDOAS SCHOLASTICA CONTINENS Diagrammen Typicam Artium: Grammatices (Latinae. Graecae.) Logices. Rhetorices. Astronomicas. Ethices. Physices. Metaphysices, seu Ontologiae. Ex Praestantium hujus temporis virorum lucubrationibus, Pro Doctrinae & virtutum studiosa juventute: CONFECTA; Iacobo Lorhardo, Gymnasij Sangallensis Rectore, & in Ecclesia Christi servo: APUD GEORGIUM STRAUB Sangalli: ANNO. 1606. It is reproduced on Raul Corazzon's deservedly award-winning website *Ontology. A resource guide for philosophers* (<http://www.formalontology.it>).
2. It seems that John McCarthy is to be credited not only for inventing the term 'Artificial Intelligence' (on the occasion of the Dartmouth conference 1955), but also for introducing *en passant* the term 'ontology' in AI when he wrote in the context of a discussion of the Missionaries and Cannibals puzzle: "Using circumscription requires that common sense knowledge be expressed in a form that says a boat can be used to cross rivers unless there is something that prevents its use. In particular, it looks like we must introduce into our *ontology* (the things that exist) a category that includes *something wrong with a boat* or a category that includes *something that may prevent its use*." (McCarthy 1980: 33–34)
3. One cannot exclude, however, that such a second metonymic extension will occur in the near future. Compare the analogously structured noun *psychology*, which already has all three readings. The main entry of the Merriam-Webster Online Dictionary (<http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=psychology&x=21&y=16>) reflects this, although in a somewhat unsystematic way: "**1** : the science of mind and behavior **2 a** : the mental or behavioral characteristics of an individual or group **b** : the study of mind and behavior in relation to a particular field of knowledge or activ-

- ity 3 : a treatise on psychology”. 2 a is the object-level reading, 3 is the meta-level reading, and 1 and 2 b represent the original meta-meta-level reading.
4. This is why publications like “Evaluating Ontological Decisions with OntoClean” (Guarino and Welty 2002) with the nice subtitle “Explosing [sic! Presumably: exploring] common misuses of the subsumption relationship and the formal basis for why they are wrong” are so important for the field.
 5. Cf. also Zaefferer (this vol.).
 6. Namely the thoughts or perceptions of those who enact the institution. The fact that (a) only the thoughts or perceptions of the institution-enacting agents, and (b) not even all of their thoughts and perceptions but only some of them are relevant for the existence of the institution in question, points to a possible revision of the notion of realism that obviates the recourse to partial non-realism even for the case of an observer who essentially co-enacts a given (for instance two-person) institution: If we redefine realism as the thesis that the objects, properties and relations the world contains exist independently of an observer’s reflecting thoughts about or perceptions of them, then I can be a realist for instance about money if I reflectingly think that this is just a disk-shaped piece of metal, but at the same time pragmatically think that this is a quarter and use it for payment, thereby enacting the institution of money.
 7. Note that the absence of commonly agreed knowledge in the objectivist sense ascribed in Section 2.4.2. of this article to *open environments* like the Semantic Web does not imply a non-realist position, it rather characterizes an epistemological situation.
 8. The point of the experiment is that as a side-effect of their specific ontology (there is no category for ‘obstacle ahead’) the Didabots clean an area cluttered with Styrofoam blocks by pushing them into clusters (Pfeifer 2000: 3), illustrating thus that an extremely simple configuration can result in complex (and useful) behavior, but this aspect cannot be elaborated on in our context.
 9. Or more precisely a cell and anatomy ontology of this roundworm (Lee and Sternberg 2003). An excerpt from the abstract: “We are endowed with a rich knowledge about *Caenorhabditis elegans*. . . . To make the information more accessible to sophisticated queries and automated retrieval systems, WormBase has begun to construct a *C. elegans* cell and anatomy ontology.”
 10. It does of course make sense to compile a corresponding database. But exactly such a database would nicely illustrate the difference between an ontological (and therefore generic) database, and a database for a specific entity token. Cf., e.g., Nirenburg and Raskin (2001: 15–16): “The following components in an agent’s model are relevant for its language processing ability: . . . Knowledge about the world, which we find useful to subdivide into:
 - an ontology, which contains knowledge about types of things (objects, processes, properties, intentions) in the world; and
 - a fact database, an episodic memory module containing knowledge about instances (tokens) of the above types and about their combinations”.
 11. “The success of a categorization can be measured by the degrees of prediction and control which the categories produced afford other scientists. Good theories are built upon successful categorizations of nature.” (Koepsell 1999: 217)

12. In the context of a talk given at the University of Munich on July 18, 1983, entitled *A Chapter of English Metaphysics*, according to the notes taken by Dietmar Zaefferer.
13. It is of course a simplification to speak of the concept of WORD, since there are several concepts we could refer to in this way, both in everyday language and in linguistic terminology, but for the present purposes these differences do not matter.

References

- Baader, Franz, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider (eds.)
2003 *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge: Cambridge University Press.
- Bach, Emmon
1986 Natural language metaphysics. In *Logic, Methodology, and Philosophy of Science VII*, Ruth Barcan Marcus, Georg J. W. Dorn and Paul Weingartner (eds.), 573–595. Amsterdam: North Holland.
1989 *Informal Lectures on Formal Semantics*. Albany, NY: State University of New York Press.
- Bailin, Sidney C., and Walt Truskowski
2003 Ontology negotiation: How agents can really get to know each other. In *Proceedings of the First International Workshop on Radical Agent Concepts (WRAC 2002)*, Walt Truskowski, Christopher Rouff, and Michael G. Hinchey (eds.), 320–334. (Lecture Notes in Computer Science 2564.) Berlin: Springer.
- Barwise, Jon, and John Perry
1983 *Situations and Attitudes*. Cambridge, MA: MIT Press.
- Bateman, John A.
this vol. Linguistic interaction and ontological mediation.
- Berkeley, George
1999 Reprint. A Treatise Concerning the Principles of Human Knowledge. In *Principles of Human Knowledge and Three Dialogues*, Howard Robinson (ed.). Oxford: Oxford University Press. Original edition, Dublin: Aaron Rhames/Jeremy Peptat, 1710.
- Bozsak, Erol, Marc Ehrig, Siegfried Handschuh, Andreas Hotho, Alexander Maedche, Boris Motik, Daniel Oberle, Christoph Schmitz, Steffen Staab, Ljiljana Stojanovic, Nenad Stojanovic, Rudi Studer, Gerd Stumme, York Sure, Julien Tane, Raphael Volz, and Valentin Zacharias
2002 KAON – Towards a large scale Semantic Web. In *Proceedings of the Third International Conference on E-Commerce and Web Technologies (EC-Web)*, Erol Bozsak, A. Min Tjoa, and Gerald Quirchmayr (eds.), 304–313. (Lecture Notes in Computer Science 2455.) Berlin: Springer.

- Buchanan, Bruce G., and Edward H. Shortliffe (eds.)
 1984 *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.
- Casati, Roberto, and Achille C. Varzi
 1995 *Holes and Other Superficialities*. Cambridge, MA: MIT Press.
- Chalupsky, Hans
 2000 OntoMorph: A translation system for symbolic knowledge. In *Proceedings 7th International Conference on Principles of Knowledge Representation and Reasoning (KR'2000)*, Fausto Giunchiglia and Bart Selman (eds.), 273–284. Breckenridge, CO: Morgan Kaufmann.
- Chaudhri, Vinay, Adam Farquhar, Richard Fikes, Peter Karp, and James Rice
 1998 OKBC: A programmatic foundation for knowledge base interoperability. In *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI'98)*, 600–607.
- Cocchiarella, Nino B.
 2001 Logic and ontology. *Axiomathes* 12 (1–2): 117–150.
- Cruse, D. Alan
 1986 *Lexical Semantics*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.
- Davidson, Donald
 1967 The logical form of action sentences. In *The Logic of Decision and Action*, Nicholas Rescher (ed.), 81–95. Pittsburgh, PA: University of Pittsburgh Press.
- Decker, Stefan, Michael Erdmann, Dieter Fensel, and Rudi Studer
 1999 Ontobroker: Ontology based access to distributed and semi-structured information. In *Semantic Issues in Multimedia Systems. Proceedings of DS-8*, Robert Meersman, Zahir Tari, and Scott M. Stevens (eds.), 351–369. Boston: Kluwer.
- Dou, Dejing, Drew McDermott, and Peishen Qi
 2002 Ontology translation by ontology merging and automated reasoning. In *Proceedings EKAW 2002 Workshop on Ontologies for Multi-Agent Systems*, Rose Dieng (ed.), 3–18.
- Elst, Ludger van, Andreas Abecker, Ansgar Bernardi, Andreas Lauer, Heiko Maus, and Sven Schwarz
 2004 An agent-based framework for distributed organizational memories. In *Proceedings Coordination and Agent Technology in Value Networks, Multikonferenz Wirtschaftsinformatik (MKWI-2004)*, Martin Bichler, Carsten Holtmann, Stefan Kirn, Jörg P. Müller, and Christof Weinhardt (eds.), 181–196, Berlin: GITO-Verlag.
- Ereshefsky, Marc
 2002 Species. In *The Stanford Encyclopedia of Philosophy* (Fall 2002 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/fall2002/entries/species/>.

- Eriksson, Henrik, Raymond W. Ferguson, Yuval Shahar, and Mark A. Musen
 1999 Automatic generation of ontology editors. In *Proceedings of the Twelfth Banff Knowledge Acquisition for Knowledge-based Systems Workshop*, Mark A. Musen and Brian R. Gaines (eds.), Banff, Alberta, Canada. Technical report, University of Calgary/Stanford University.
- Farrar, Scott
 this vol. Using ‘Ontolinguistics’ for language description.
- Fellbaum, Christiane
 this vol. The ontological loneliness of verb phrase idioms.
- Fellbaum, Christiane (ed.)
 1998 *WordNet: An Electronic Lexical Database*. (Language, Speech, and Communication.) Cambridge, MA/London: MIT Press.
- Fischer, Felix, and Matthias Nickles
 to appear Computational opinions. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI-06)*, Gerhard Brewka (ed.). Amsterdam: IOS Press.
- Froehner, Tina, Matthias Nickles, and Gerhard Weiss
 2004 Towards modeling the social layer of emergent knowledge using open ontologies. In *Proceedings of The ECAI 2004 Workshop on Agent-Mediated Knowledge Management (AMKM-04)*, Andreas Abecker, Ludger van Elst, and Virginia Dignum (eds.), 10–19.
- Gangemi, Aldo, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider
 2002a Sweetening ontologies with DOLCE. In *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management*, 166–181. Berlin/Heidelberg: Springer.
- Gangemi, Aldo, Nicola Guarino, Alessandro Oltramari, and Stefano Borgo
 2002b Cleaning-up WordNet’s top-level. In *Proceedings of the First Global WordNet Conference*. Mysore: Central Institute of Indian Languages.
- Genesereth, Michael
 1991 Knowledge Interchange Format. In *Proceedings of the Second International Conference on the Principles of Knowledge Representation and Reasoning*, James F. Allen, Richard Fikes, and Erik Sandewall (eds.), 238–249. San Mateo, CA: Morgan Kaufmann.
- Giugno, Rosalba, and Thomas Lukasiewicz
 2002 P-shoq(d): A probabilistic extension of shoq(d) for probabilistic ontologies in the semantic web. In *JELIA’02: Proceedings of the European Conference on Logics in Artificial Intelligence*, Sergio Flesca, Sergio Greco, Nicola Leone, and Giovambattista Ianni (eds.), 86–97, Berlin: Springer.
- Goddard, Cliff
 this vol. Semantic primes and conceptual ontology.

- Golbeck, Jennifer, Bijan Parsia, and James A. Hendler
 2003 Trust Networks on the Semantic Web. In *Proceedings of the Seventh International Workshop CIA-2003 on Cooperative Information Agents*, Matthias Klusch, Sascha Ossowski, Andrea Omicini, and Heimo Laamanen (eds.), 238–249. (Lecture Notes in Computer Science 2782.) Berlin: Springer.
- Gruber, Thomas R.
 1993 A translation approach to portable ontology specification. *Knowledge Acquisition* 5: 199–220.
- Grüninger, Michael, and Christopher Menzel
 2003 Process Specification Language: Principles and applications. *AI Magazine* 24 (3) (Fall 2003): 63–74.
- Guarino, Nicola, and Christopher Welty
 2002 Evaluating ontological decisions with OntoClean. *Communications of the ACM* 45 (2): 61–65.
- Haider, Hubert
 1993 *Deutsche Syntax – generativ*. Tübingen: Narr.
- Heflin, Jeff, and James A. Hendler
 2000 Dynamic ontologies on the web. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence (AAAI-2000)*, 443–449. Menlo Park, CA: AAAI/MIT Press.
- Hellwig, Birgit
 this vol. Postural categories and the classification of nominal concepts: A case study of Goemai.
- Hofweber, Thomas
 2005 A puzzle about ontology. *Nous* 39 (2): 256–283.
- Huang, Zhisheng, Frank van Harmelen, and Annette ten Teije
 2005 Reasoning with inconsistent ontologies. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-2005)*, Leslie Pack Kaelbling, and Alessandro Saffiotti (eds.), 454–459. Edinburgh: Professional Book Center.
- Imai, Mutsumi, and Dedre Gentner
 1997 A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62: 169–200.
- Khrentzos, Drew
 2004 Semantic challenges to realism. In *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2004/entries/realism-sem-challenge/>.
- Koepsell, David R.
 1999 Introduction to applied ontology: The philosophical analyses of everyday objects. *American Journal of Economics & Sociology* 58 (2): 217–220.

- Lang, Ewald
1984 *The Semantics of Coordination*. Amsterdam: John Benjamins.
- Lee, Raymond Y.N., and Paul W. Sternberg
2003 Building a cell and anatomy ontology of *Caenorhabditis elegans*. *Comparative and Functional Genomics* 4 (1): 121–126.
- Lenat, Douglas B.
1995 CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM* 38 (11): 33–38.
- Link, Godehard
1983 The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, Use, and the Interpretation of language*, Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.), 303–323. Berlin/New York: Walter de Gruyter.
- Luger, George F., and William A. Stubblefield
1993 *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Redwood City, CA: Benjamin/Cummings.
- Martin, Philippe, and Martin Eklund
2001 Large-scale cooperatively-built heterogeneous KBs. In *Proceedings of the 9th International Conference on Conceptual Structures*, Gerd Stumme and Guy W. Mineau (eds.). (Lecture Notes in Artificial Intelligence 2120.) Berlin: Springer Verlag.
- McCarthy, John
1980 Circumscription – A form of non-monotonic reasoning. *Artificial Intelligence* 13: 27–39.
1987 Generality in Artificial Intelligence. *Communications of ACM* 30 (12): 1030–1035.
- McCarthy, John, and Pat Hayes
1969 Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence* 4: 463–502.
- McGuinness, Deborah, and Frank van Harmelen
2004 OWL web ontology language: Overview. (W3C Recommendation 10, February 2004.) <http://www.w3.org/TR/owl-features/>.
- Mel'cuk, Igor
1998 Collocations and lexical functions. In *Phraseology. Theory, Analysis, and Applications*, Anthony P. Cowie (ed.), 23–53. Oxford: Clarendon Press.
- Melle, William van
1982 *System Aids in Constructing Consultation Programs: EMYCIN*. Ann Arbor, MI: UMI Research Press.
- Mena, Eduardo, Arantza Illarramendi, Vipul Kashyap, and Amit Sheth
2000 OBSERVER: An approach for query processing in global information systems based on interoperation across pre-existing ontologies. *On Distributed and Parallel Databases – An International Journal* 8 (2): 223–271.

- Metzinger, Thomas, and Vittorio Gallese
 this vol. The emergence of a shared action ontology: Building blocks for a theory.
- Miller, Alexander
 2005 Realism. In *The Stanford Encyclopedia of Philosophy* (Fall 2005 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/fall2005/entries/realism/>.
- Miller, George A.
 1995 WordNet: A lexical database for English. *Communications of the ACM* 38 (11): 39–41.
- Misskof, Michele, Roberto Navigli, and Paola Velardi
 2002 Integrated approach to web ontology learning and engineering. *IEEE Computer* 35 (11): 60–63.
- Montague, Richard
 1973 The proper treatment of quantification in ordinary English. In *Approaches to Natural Language*, K. Jaakko J. Hintikka, Julius M. E. Moravcsik, and Patrick Suppes (eds.), 221–242. Dordrecht: Reidel.
- Motik, Boris
 2005 On the properties of metamodeling in OWL. In *Proceedings of the 4th International Semantic Web Conference (ISWC 2005)*, Yolanda Gil, Enrico Motta, V. Richard Benjamins, and Mark A. Musen (eds.), 548–562. (Lecture Notes in Computer Science 3729.) Berlin: Springer.
- Müller, Heinz-Jürgen, and Rose Dieng (eds.)
 2002 *Computational Conflicts*. Berlin: Springer.
- Nadathur, Gopalan, and Dale Miller
 1998 Higher-order logic programming. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 5, Dov M. Gabbay, Christopher J. Hogger, and John A. Robinson (eds.), 499–590. Oxford: Oxford University Press.
- Nickles, Matthias, Tina Froehner, Ruth Cobos, and Gerhard Weiss
 2005 Multi-source knowledge bases and ontologies with multiple individual and social viewpoints. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, Andrzej Skowron, Rakesh Agrawal, Michael Luck, Takahira Yamaguchi, Pierre Morizet-Mahoudeaux, Jiming Liu, and Ning Zhong (eds.), 62–65. Los Alamitos, CA: IEEE Computer Society Press.
- Niles, Ian, and Adam Pease
 2001 Toward a Standard Upper Ontology. In *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)*, Christopher Welty and Barry Smith (eds.), 2–9. New York: ACM Press.
 2003 Linking lexicons and ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the IEEE International*

Conference on Information and Knowledge Engineering, Nazli Goharian (ed.), 412–416. Las Vegas, NV: CSREA Press.

Nirenburg, Sergei, and Victor Raskin

2004 *Ontological Semantics*. Cambridge, MA: MIT Press.

Nodine, Marian, Jerry Fowler, and Brad Perry

1999 An overview of active information gathering in InfoSleuth. In *Proceedings of the Second International Symposium on Cooperative Database Systems for Advanced Applications (CODAS'99)*. Berlin: Springer.

Object Management Group

1997–2006 UML™ Resource Page. Available online at <http://www.uml.org/> (accessed 30 May 2006).

Pease, Adam

2003 The Sigma ontology development environment. In *Working Notes of the IJCAI-2003 Workshop on Ontology and Distributed Systems*, Fausto Giunchiglia, Asucion Gomez-Perez, Adam Pease, Heiner Stuckenschmidt, York Sure, and Steven Willmott (eds.). (CEUR Workshop Proceeding Series 71.) <http://ceur-ws.org/Vol-71/>.

this vol. Formal representation of concepts: The Suggested Upper Merged Ontology and its use in linguistics.

Pease, Adam, and Christiane Fellbaum

2004 Language to logic translation with PhraseBank. In *Proceedings of the Second International WordNet Conference (GWC 2004)*, Petr Sojka, Karel Pala, Pavel Smrz, Christiane Fellbaum, and Piek Vossen (eds.), 187–192. Brno: Masaryk University.

Pease, Adam, Raymond A. Liuzzi, and David Gunning

2001 Knowledge bases. In *Encyclopedia of Software Engineering*, John J. Marciniak (ed.). 2nd ed. New York: John Wiley and Sons.

Pfeifer, Rolf

2000 On the role of morphology and materials in adaptive behavior. In *From Animals to Animats. Proceedings of the 6th International Conference on Simulation of Adaptive Behavior*, Jean-Arcady Meyer, Alain Berthoz, Dario Floreano, Herbert L. Roitblat, and Stewart W. Wilson (eds.), 23–32. Cambridge, MA: MIT Press.

Ray, Steven

2004 NIST's semantic approach to standards and interoperability. Unpublished presentation.

Richardson, Matthew, Rakesh Agrawal, and Pedro Domingos

2003 Trust management for the Semantic Web. In *Proceedings of the International Semantic Web Conference (ISWC-03)*, Dieter Fensel, Kattia Sycara, and John Mylopoulos (eds.), 351–368. (Lecture Notes in Computer Science 2870.) Berlin: Springer.

- Schalley, Andrea C.
2004 *Cognitive Modeling and Verbal Semantics. A Representational Framework Based on UML.* (Trends in Linguistics. Studies and Monographs 154.) Berlin/New York: Mouton de Gruyter.
- Schalley, Andrea C., and Dietmar Zaefferer
this vol. Ontolinguistics – An outline.
- Searle, John R.
2005 What is an institution? *Journal of Institutional Economics* 1 (1): 1–22.
- Staab, Steffen, Alexander Mädche, Frank Nack, Simone Santini, and Luc Steels
2002 Emergent Semantics. *IEEE Intelligent Systems, Trends & Controversies* 17 (1): 78–86.
- Staab, Steffen, and Rudi Studer (eds.)
2003 *Handbook on Ontologies in Information Systems.* Berlin: Springer.
- Talmy, Leonard
this vol. The representation of spatial structure in spoken and signed language: A neural model.
- Tamma, Valentina A. M.
2002 *An ontology model supporting multiple ontologies for knowledge sharing.* Ph.D. diss., Department of Computer Science, University of Liverpool.
- Yergeau, François, Tim Bray, Jean Paoli, C. Michael Sperberg-McQueen, and Eve Maler
2004 Extensible Markup Language (XML) 1.0 (Third Edition). (W3C Recommendation 4th February 2004.) <http://www.w3.org/TR/2004/REC-xml-20040204/>
- Zaefferer, Dietmar
2002 Polysemy, polyvalence, and linking mismatches. The concept of RAIN and its codings in English, German, Italian, and Spanish. *DELTA – Documentação de Estudos em Lingüística Teórica e Aplicada* 18 (spe.): 27–56. Special Issue: Polysemy.
this vol. Language as mind sharing device: Mental and linguistic concepts in a general ontology of everyday life.