DISTINCTIONS—A COMMON BASE FOR A TAXONOMIC CALCULUS FOR OBJECTS AND ACTIONS

Andrew U. Frank,
Institute for Geoinformation and Cartography,
Technical University Vienna
{frank@geoinfo.tuwien.ac.at}

Note to editor: figures and pictures for a publication will be improved.

1 Introduction

Just as using concepts and predictive models to understand the world is a fundamentally human activity, so is organizing these concepts into systems, which aid understanding of concepts and the world. Ontologists investigate formalizations of these concept organizations for various applications, especially data and operations in information systems.

The major components of ontologies are taxonomies of objects, representing the *is_a* relation. In the18th century Linnaeus showed that taxonomies are an excellent tool to structure knowledge; his taxonomy of plants has advanced biological research enormously and is still in use today. The application of similar principles in library science to organize books by Melville Dewey in the 19th century led to a number of taxonomies of knowledge in general. Taxonomies of the objects in a universe of discourse reveal much of the conceptualization of the subset of reality considered, often making clear its prejudices and blind spots.

This article applies and generalizes taxonomic methods that have been successful for classifying the static aspects of our conceptualizations to a classification of the dynamic aspects of reality, i.e., a classification of the actions that change the world. The calculus for taxonomies, which classify objects and actions, addresses open problems with taxonomies:

- Ontologists take the conceptual structure represented in language as a starting point, despite empirical evidence that mental concepts have a finer and possibly different structure than the words of a natural language (Fodor 1984; Fodor 1995; Jackendoff 1996); I start with distinctions between objects and construct taxa as fine-grained as they are required within a situation.
- Most taxonomies organize classifications of objects. I demonstrate methods for constructing taxonomies of actions as well as objects using the same distinctions and the same construction principles.
- Taxonomies are often constructed hierarchically, where every taxon is linked to exactly one taxon above it. This restriction leads to incompatible taxonomies for closely related fields and makes the integration of knowledge organized in this form difficult (Guarino and Welty

2000). Merging the taxonomic lattice constructed from different sets of distinctions is reduced to the identification of the common distinctions. This reduces the integration task even for very large taxonomies to a manageable problem.

The approach focuses on taxonomies for physical reality; the application of the same methods to socially constructed reality and subjective mental states is left for future work. The goal is a calculus for taxonomies and therefore the focus is on taxonomic classes, not the individuals as in some other ontology research (BFO (Grenon, Smith et al. 2004), DOLCE (Gangemi, Guarino et al. 2002)). At the same time, the interest is in mental concepts, not the concepts expressed in linguistic units, e.g., words of the lexicon of a natural language, as in WordNet (Fellbaum 1998).

The approach here starts with *distinctions* between objects and applies them to the classification of actions. Distinctions are experienced, meaningful differences between objects; objects that are similar with respect to a distinction are classified in a taxon; this relates the taxa to differences between the objects, and connects taxonomies to affordances (Gibson 1986; Raubal 2000). Which distinctions are meaningful is context dependent and varies with situation, person, intended goals, etc. The often suggested approach to identify the nouns and verbs in a natural language description of the field (Parnas and Share 1978; Kuhn 1994) is a viable method to identify the relevant distinctions.

A taxon is a set of distinction values; for example, animals are physical objects that are alive. The taxa are partially ordered by a subset relation. All possible taxa produced from a fixed set of distinctions form a lattice with *join* and *meet* operations. The *join* operation gives the least common superclass and the *meet* operation the most general common subclass for two taxa. The construction of a taxonomic lattice from a set of distinctions gives a taxonomy that is in a certain sense of a uniform level of detail and constructs ontologies that are consistent in what differentiations are included and what is excluded.

The article shows how much of ontological reasoning can be achieved with a calculus on taxonomies and does not require complex, model-based logical deductions. The formalization as a lattice structure leads to an algorithmic approach for the integration of different taxonomies; it is not based on a similarity measure that may not produce correct results for taxonomies built from substantially different viewpoints.

In the next section a running example is introduced. Section 3 then clarifies terminology. Section 4 explains how taxonomies for objects are constructed from distinctions. In section 5 general properties of taxonomic lattices are described and section 6 gives a method of classifying relations. Section 7 starts with the ontological commitments for actions and constructs taxonomies for them. Section 8 sketches the application of the described taxonomies of mental concepts to communication, which is mostly left to future work. The concluding section reviews the solution.

2 Running Example

2.1 Single action

Picture 1 was taken on 28th of May 2006 at 9:55 in my kitchen in Geras (Austria). It shows Mary cutting a piece of bread from a loaf of bread with a bread knife from IKEA. This picture shows an action; the world—in particular the loaf of bread—is changed by this action (compare picture 1 before and 2 after the cutting was completed), but in smaller ways the person involved, the knife, and the table have also slightly changed. The changes in the mass of the loaf are most noticeable, but there was also the movement of a part of the loaf to a new position on the table, the knife has been dulled, etc.

Picture 1: An Action

Picture 2: The effect of an action

In the action shown, the following objects are involved: a loaf of bread, a knife, a person, and a table. These objects have each an identity; for example, the bread knife is the individual I have bought from IKEA a while ago; it has the same form as countless other similar knives manufactured by the same company and distributed all over the world, but this individual is different from any other similar copy; it will always be at a different location than any of the other similar bread knives. The loaf of bread is made from flour, it is solid, has a specific volume and a mass (volume and mass are roughly proportional). For purposes of this article, the changes caused by the action shown are only in the volume and mass of the loaf of bread and the appearance of a new piece (Picture 2). The minimal changes in knife, person, and table are not considered; the qualities of these objects remain the same.

2.2 Extended example

Consider the following sequence of actions: Peter walks to the store in the morning of the same day at 7:35 and buys milk, butter, and bread. He walks back with the groceries in his basket. He puts bread on the table and the milk and butter in the fridge in the kitchen. Mary prepares the table for breakfast with plates, cups, and knives (Picture 3), which she takes from a cupboard and a drawer. She cuts a piece of bread. She gets butter and cheese from the fridge and puts it on the table. She puts butter on the piece of bread, which she then eats.

Picture 3: The kitchen table set for breakfast

Picture 4: Five similar knives

We could have observed the scene with a camera and produced a video-clip that shows the sequence of actions, as they occurred on this day, at a specific time and a specific location. Peter and Mary report this morning's actions to David later; differences in *actions* and *accounts* of actions will be touched upon briefly in section 8.

2.3 Universe of discourse

The description gives a list of objects that is considered in the universe of discourse constituted by the running example: bread, table, knife, Mary, Peter, butter, cheese, milk, etc. These nouns describe classes of material objects (count nouns), sometimes materials (mass nouns) but they are used in the context to describe specific individuals of these object classes or materials, even though it does not matter which one of an equal set of individuals is used (La Palme Reyes, Macnamara et al. 2002). It is (usually) inconsequential which one of the 5 knives from the drawer Mary uses to butter her bread (Picture 4), but the 5 knives nevertheless are distinct individuals, each having a distinct location and by very accurate measurement, a distinct (but very similar) weight, form, material, etc.

The actions described are: the walking to the store, gathering objects there and carrying them back, putting things into storage and moving them out of storage, dividing bread and putting butter on the piece of bread. Each action occurs at a specific time and location with specific objects.

Actions occurring in the world are always affecting and involving specific individual objects and occur at specific time and location; we say the objects inhere in the action. *Reality is fully detailed*. Only the account given above is a generalization, reducing the nearly infinite amount of detail to what is necessary for the reader to reconstruct the actions. The interpretation of the generic count and mass nouns is to pick the unique individual described by the context.

3 Terminology

As far as possible I will follow the terminology used in the WonderWeb project as their ontological commitments are clearly documented and to a large degree similar to the view used here (Masolo, Borgo et al. 2003). I will use the word *concept* for human mental constructs and data or *representations* for external signs that are related to concepts.

I will use *object* or *individual* to describe a single, particular object (e.g., the loaf of bread in Picture 1) and for particular changes in the world I will use the term *actions*. Examples for individuals in formulae will be written with lower case letters (and often an appended number). Individuals have qualities, which can be observed and are related to values in a quality space (Gärdenfors 2000) "within a certain ontology, we assume that these qualities belong to a finite set of quality types" (Masolo, Borgo et al. 2003, 16). For the purposes of this article, we have *properties* as universals, which are functions that map from an individual to a *property value*; property values represent the qualities of the individual. Distinctions are particular types of properties; they are capitalized.

Humans communicate about situations and describe actions and individuals involved in them. These *accounts* of actions are less detailed than a mental model and report only what seems relevant to the communicating person; they are dependent on the knowledge a person has of a situation, what she considers important and also what the person intends to communicate or, possibly, to hide, about her knowledge of a situation.

4 Taxonomies for Objects

Humans classify their conceptualization of objects into sets where all objects in the set have some similarity and are distinct from other individuals, e.g., Humans, Dogs, Tables. The distinction identifies aspects that are important for the situation and how humans interact with the individuals involved. The same set of distinctions is not always relevant and the classification is therefore context dependent. The same individual can be classified differently in different situations. Consider what is called a chair in different situations (Picture 5).

Picture 5: Different types of chairs

Natural languages use nouns to describe objects and verbs to represent actions. Ignoring polysemy, nouns represent intensional sets of objects that are similar in some respect and this similarity is important in the context the noun is used. Nouns are most often used to refer to a single individual of the class named; one speaks about the loaf of bread, the bread knife, meaning the single individual present in the context. Verbs pick out similarity in a large set of actions and group actions with respect to a property relevant in the context. I will show in section 7 that differences in operations are directly related to the properties of the objects involved; Gibson introduced the term affordance to describe the relation between qualities of the environment and the objects in it and the operations that are possible (Gibson 1986).

4.1 Distinctions

The focus in this work is on differences in the values of properties of an object humans distinguish and use to distinguish it from others. Distinctions are used to form a conceptual set of objects. Distinctions are properties that map from objects, situations, or actions to a small set of values, often just the Boolean true or false.

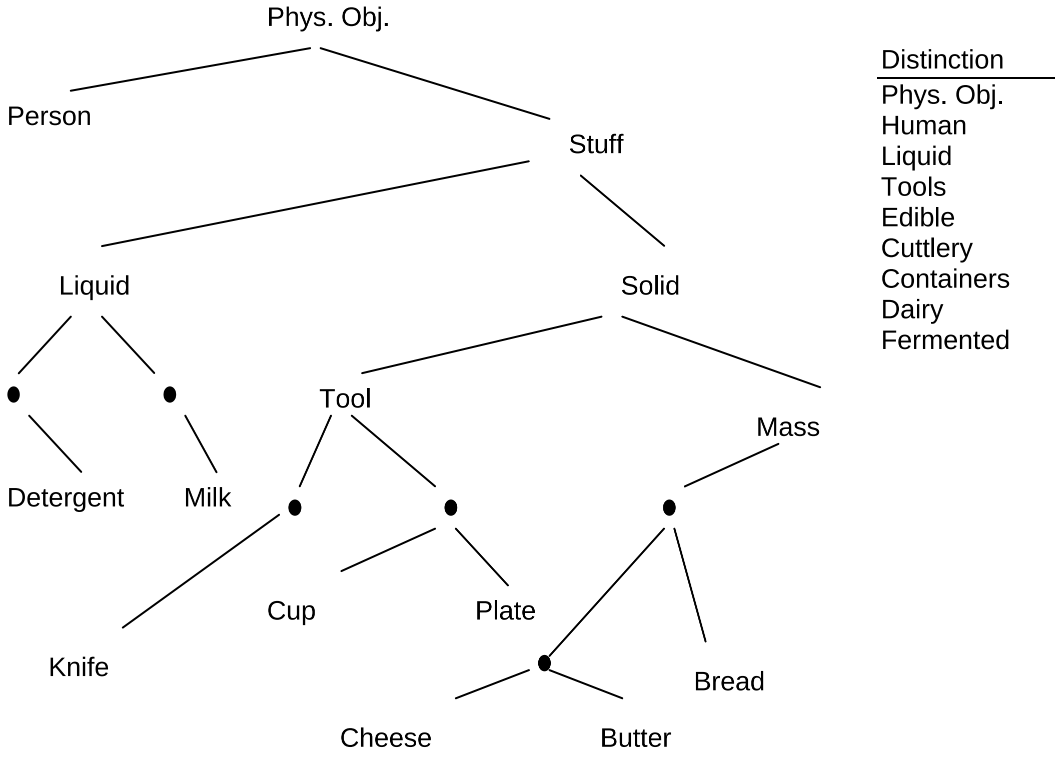

Figure 1: A taxonomy for part of the running example

The distinctions in the example are, for example: *Human*, *PhysicalObject*, and *Liquid*; the complete list of distinctions considered here is found in the right margin of Figure 1. The distinctions relate to experiential reality: observations or actions performed with such objects; natural languages nouns and verbs describe intensional sets of objects or actions that are similar in a property.

If ordinary *properties p* are seen as functions from objects to property values from a continuous domain ($p :: o \text{ -> } v$), then distinctions *d* are a composition of a property function and a mapping *eq* from property values *v* to an equivalence classes *w*, reducing the domain to a small number of discrete values:

$$eq :: v \text{ -> } w, \;\; d : o \;\text{ -> } w$$
$$d = p \;.\; eq$$

To simplify the discussion in this article, only Boolean distinctions are used and the names are selected such that the name relates to the taxon for which the distinction obtains (has the value true); I will show in subsection 5.7 that this is not a restrictive assumption and the general case of distinctions that allow several values can be mapped to several Boolean distinctions without loss.

4.2 Taxon

A taxon is a set of objects such that all objects have a specific value for a distinction. All objects in the taxon *Solid* have for the distinction *Liquid* the value *false*. A function *d'* maps from a taxon to a value for a distinction (*d' :: t -> w*) ; this corresponds to the ordinary language statement "Dogs eat meat", meaning "if x is a dog then x eats meat" (properly speaking, the taxon *Dog* is an abstract concept that cannot "eat"). The interpretation of a taxon t_i is therefore:

$$o \in t_i => d'_i (t_i) = d_i (o).$$

If an individual *o* is member of a taxon *t* then the value of the distinction d_i for this object $d_i (o)$ is the value of the distinction for the taxon $d'_i (t_i)$; this entails that the value for a distinction is the same for all objects in the respective taxon (Figure 2):

$$o_1, o_2 \in t_i <=> d_i (o_1) == d_i (o_2).$$

A taxon is essentially an equivalence class for the individuals so classified (taxed?).

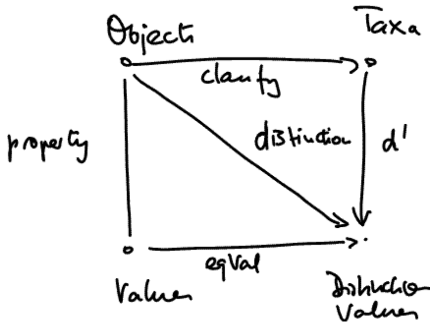

Figure 2: Category diagram showing the relation between objects and taxa (*p* is the property function, *d* and *d'* are the distinctions applied to objects and taxa respectively)

4.3 Intensional definition of taxa

Sets of individuals can be defined intensionally, i.e., defined by a general rule to decide if an individual is member of the set, or extensionally, i.e., by enumeration of the members. A taxon is constructed intensionally as a set of individuals such that they have the same value for a specific property and distinct values for individuals that are members of other taxa. If a taxon is described extensionally by pointing to a set of individuals, e.g., by stating "these are all chairs!" (Picture 5), then it is left to the hearer to understand what aspect of similarity is assumed and construct the intension in the taxon.

4.4 Characterization of taxa

A taxon is characterized by the distinctions and the values obtained; for example the taxon *Milk* can be described as the set *{PhysObj+, Human-, Liquid+, Edible+, Tool-},* where a plus or minus after the distinction name indicates whether the value for the distinction is *true* or *false*. Note that a taxon is described by the set of the distinction values, the order in which they are listed is not relevant.

This notation corresponds to the use of positive and negative tags by Gruber (2005). In (Frank to appear 2006) I grouped the affirmed and the negated distinctions in two sets that characterized a taxon; this may be advantageous for certain implementations and is logically equivalent.

4.5 Rule of full distinction

For a universe of discourse, a *rule of full distinction* applies: There must be a sufficient number of distinctions to separate all taxa; this requires for n taxa at least $\log_3 n$ and at most n distinctions. Figure 1 shows how 8 distinctions define (at least) 14 taxa.

4.6 Named taxa

The distinction and rules construct all possible taxa in a taxonomic lattice. Some of these taxa are important in the universe of discourse and are singled out; ontologists typically assign names to them as a form of quick reference. We will in the following use the names for the taxa that are shown in Figure 1. Note that names for taxa are just shorthand for the characteristic set of distinctions and for the calculus names are always replaced by the corresponding set of distinction values; e.g.,

Bread = {PhysObj+, Human-, Liquid-, Edible+, Tool-, Dairy-}.

5 Calculus of Taxonomy

From a set of distinctions $D = \{d_i\}$ a taxonomic lattice is constructed. A lattice is a partially ordered structure whose nonempty finite subsets all have a supremum (called *join*) and an infimum (called *meet*). A lattice can be defined equivalently as an algebraic structure with two binary operations *join* and *meet* that are idempotent, commutative, associative and for which an absorption law $a \wedge (a \vee b) = a$ and $a \vee (a \wedge b) = a$ is valid. A lattice has a greatest element (*top*) and a least element (*bottom*). For a taxonomic lattice, the *top* is interpreted as the taxon that includes everything, and *bottom* is the taxon that includes nothing.

For each Boolean distinction, a taxon can have one of four values: the value of the distinction is *irrelevant* (the taxon is indifferent to this value), the distinction *applies* (is true) or is *negated* (is false), or the distinction applies and does not apply, which is a *contradiction*. These four values can be ordered such that we obtain the lattice shown in Figure 3 (where elements above are considered >). The logical interpretation uses the four valued logic:

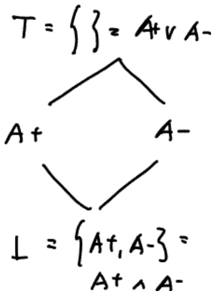

Figure 3 The lattice T_1 for a single distinction

Belnap has proposed a four valued logic with the constants *{(A or not A), A, not A, (A and not A)}* where *A or not A* is interpreted as 'indifferent' (absence of knowledge) and *A and not A* as contradiction (undefined) (Belnap 1977).

5.1 Combination of multiple distinctions

The taxonomy resulting from the combination of two distinctions *A* and *B* is the product of multiple lattices T_1 with an identification of all products that contain $\perp$ (contradiction) to *bottom* and identifying $\{\{\}_A\}, \{\{\}_B\}$ with just $\{\}$. Figure 4 shows the product of $T_1 \times T_1$. It is easy to demonstrate that the identification of all products that contain *bottom* does preserve the order relation between lattice elements.

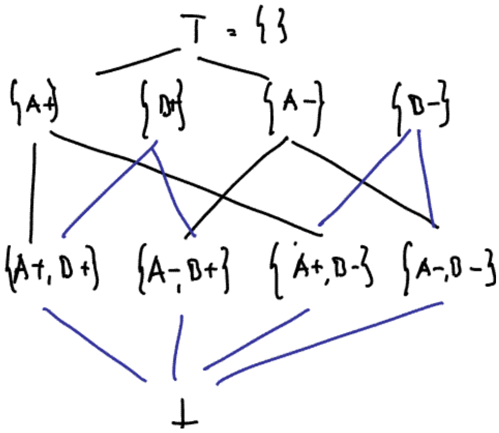

Figure 4 The lattice T_2 constructed from 2 distinctions

From any set of Boolean distinctions *D* a taxonomic lattice T_n as the product of *n* copies of T_1 can be constructed ($T_n = T_1^n$, where $n = cardinality\ D$).

The identification of all combinations that contain bottom in any of the distinctions with the single element bottom reduces the number of elements for T_2 from $4 * 4 = (2^2 * 2^2)$ to only *10*; the combinations $\{\perp, A+\}$ $(\perp, A-)$, etc. are mapped to $\perp$. In general, a taxonomic lattice with *n* distinctions has $3^n + 1$ element. The *8* distinctions shown in Figure 1 produce a taxonomic lattice that contains at most 730 taxa—only a small number are shown in Figure 1!

5.2 Order relation between taxa

A partial order relation < exists between two taxa; a taxon *S* is a subtaxon of a taxon *P* if the characterizing set of distinction values of *S* is a subset of the characterizing set of *P*. For example, the taxon *Milk* is a subtaxon of *Liquids = {PhysObj+, Human-, Liquid+}.*

It is possible to construct taxa so fine that each taxa contains exactly one individual—for example using distinction of where and when an object is present. For these finest taxa the intensional definition by distinction and the extensional definition by enumeration connect.

5.3 Interpretation of *join* and *meet*

In a lattice, a *join* and a *meet* exist for any two elements; they are defined as the intersection (respective union) of the characteristic sets (*similar* to *join* and *meet* in the lattice of the powerset of a set *D*). The interpretation of *join* is the least general taxon that has the joint properties of the given one; in Figure 1 the *join* of *Milk* and *Bread* is {*PhysObj+,Human-, Edible+*}—this taxon is not shown in the subset of the taxonomic lattice shown in Figure 1 but is included in the taxonomic lattice. *Joins* may be interesting and useful taxa—for example the *join* of *Dogs* and *Car* is a *Self-propelled Land Object*—pointing to the few, but important similarities.

Drawing a taxonomic lattice is—for any than the smallest examples – not practical. I show here sections of lattices such that the distinctions are applied in a specific order; they appear as hierarchical taxonomies (often without *top* and always without *bottom*). Note that these show only parts of a lattice and that many more taxa are virtually included such that for any two nodes a *join* (common superclass) and a *meet* (common subclass) exist—in some cases these may be *top* or *bottom*.

5.4 Dependencies between distinctions

A dependency relation exists between the distinctions: it is not meaningful to question whether an object is liquid unless it is a physical object. *Cups* must be *Solid+*, *Tools* must be *Human-*, etc. These restrictions can be used to reduce the size of a taxonomic lattice further. Rules express dependencies, e.g.:

{Human+} -> {Tool-} or
{Solid-} -> {Tool-}.

Characterizations of taxa that violate these rules are mapped to *bottom*, resulting in the taxon that has no possible object in it. This suggests an alternative representation of rules as patterns that map to bottom: *{Human+, Tool+}, {Solid-, Tool+},* etc.

These dependencies express the intended semantics of the distinctions. One could cast the semantic net for the distinctions somewhat wider, allowing tools to be human or non-human, including "employee" as a subtaxa of {*Tools*+} or understand detergent as a taxon {*Liquid*+, *Tools*+}.

5.5 Integration of taxonomies for objects

Extensive research has focused on methods integrating ontologies, especially the entity-relationship (database schema) components (Claramunt and Thériault 1996; Hornsby and Egenhofer 1997; Fonseca, Egenhofer et al. 2002). Such integration of ER diagrams is crucial for query interoperability and the integration of different databases.

Practical approaches work by standardizing vocabularies, but these approaches are limited to applications that are thematically connected. The integration of ontologies that were constructed by different applications, using necessarily different conceptualizations, is an open problem. Approaches that use structural similarity measures in the relations between the taxa will work only for taxa that are central and common for both applications—but integration is particularly important when data from two applications with little conceptual overlap must be combined; in such situations, similarity measures will be misleading.

Integration of two taxonomic lattices T_1 and T_2, which were constructed from two sets of distinctions D_1 and D_2 requires only the integration of the two sets of distinctions. It may be necessary to rename the distinctions in order to avoid that two semantically different distinctions from D_1 and D_2 have the same name and to assure that a distinction that has the same semantics in both taxonomies goes by the same name. The taxonomic lattice resulting from the union of D_1 and D_2 is the desired integrated taxonomy $T_m = T_1 \cup T_2$. The integration of the small number of distinctions using expert knowledge in this method is much more feasible than an integration of the much more numerous taxa in other methods. Working with distinctions reduces the effort from n to $log_3 n$. I also assume that the semantics of distinctions is less influenced by prototype effects (Rosch 1973), but this requires empirical observations.

Merging two taxonomies with differences in the set of distinctions may produce taxa that are extensionally the same but are differently described. The distinction between *Butter* and *Cheese* was influenced by the production process and labeled as "fermented"; an alternative taxonomy could label *Butter* as {*Fat*+}. Merging the two taxonomies should not pretend that this is the same distinction (*Fermented* could have been used as a distinction between black and green tea!), but realize that $butter_1$ and $butter_2$ are both superclasses of $butter_{m.}$.

5.6 Distinctions, which are not Boolean

Consider the distinction habitat for birds, (land) animals, and fish. Habitat has one of three values *{Land, Air, Sea}*. This is an example of a distinction that is not Boolean, but can be translated to three Boolean distinctions: *Land, Air,* and *Sea*. Birds would have values then binary distinctions *{Land-, Air+, Sea-}*. Rules exclude that an animal taxon has more than one true habitat value; *{Land+, Air+)}, {Land+, Sea+},* etc. are all mapped to *bottom*. This approach shows how to reduce all distinctions to Boolean distinctions. Note that a slightly different semantics for habitat would suggest that the taxon *Waterfowl* is *{Land+, Air+, Sea+}*, etc.

6 Taxonomy of Relations

A taxonomy for relations, like "the bread is on the table", can be constructed from distinctions following the same principles. The characteristic function for a relation maps from a pair of individuals to a Boolean value. "The bread is on the table" maps to the characteristic function *on (bread, table)* -> *True*.

A taxonomy for binary relation is thus a product of two taxonomies for objects. A relation *on,* which obtains between any type of physical objects, is a taxon described as a pair of object taxa *({PhysObj+}, {PhysObj+})*. The (spatial) relation *at,* which obtains between a physical object and a location, is the taxon *({PysObj+},{Location+}),* etc.

If a relation, e.g., containment, is only possible if the objects involved have some properties (what Gibson calls 'affordance' (Gibson 1986)), then a corresponding distinction is introduced. Such additional distinctions distinguished relations that would have otherwise the same set of distinctions. For example, between physical objects the relations *on* (support) and *in* (containment) are possible. They are differentiated by requiring that the second object can provide support or containment.

On: ({PhysObj+}, {PhysObj+, Support+})
In: ({PhysObj+}, {PhysObj+, Container+})

7 Taxonomies for a Dynamic World

The taxonomy of objects assumed that the world is conceptualized as consisting of objects, which are not necessarily not overlapping (not JEPD), and that the objects have qualities. Properties were defined as mappings from objects to property values and some particular properties were identified as distinctions and used to differentiate the taxa.

To extend this static view of the world to a dynamic world, I conceptualize the world as having a state at each moment in time. The state of the world evolves over time. Individuals (objects) are conceptual units that have a state, part of the state of the world, and actions are conceptual units that change the state of objects in time. The world is the totality of all individuals and the world state is the sum of the states of the individual at any moment in time; the world state evolves through the individuals evolution of the object states.

Actions are conceptualizations at the same level as individuals—they stand for a particular, fully determined change in the state of an object and link it to actions of other objects.

7.1 Example action:

Picture 1 shows an action that Mary conceptualizes as:

newWorldState =
cutBread_fig1 (mary, breadKnive_1, breadLoaf_2, kitchenTable_3,
time_now, location_here, currentWorldState)

and that changes the *currentWorldState* to a *newWorldState* and creates *pieceOfBread_7*. The individuals involved in the action (often named with appended identifiers) are classified in taxa. This gives a taxonomic description of the action, which compares directly with the type of a function in a typed programming language (Jensen and Wirth 1975; Peterson, Hammond et al. 1997; Peyton Jones, Hughes et al. 1999).

cutBread_1 :: Women x Knife x Bread x May2006 x Geras -> WorldState ->
WorldState

In this taxon for an action, the names of object taxa are used as shorthands for the corresponding sets of distinctions, e.g., *Women = {PhysObj+, Human+, Female+}; Knife = {PhysObj+, Tool+},* etc. *May2006* stands for any time point in May 2006 and *Geras* for all locations in the town of Geras (Austria). For the calculus with action taxa, the abbreviations are expanded to

cutBread_1 :: {PhysObj+, Human+, Female+} x {PhysObj+, Tool+, Cutlery+}x...

7.2 Generalization of actions

A partial order between taxa of actions is defined similar to and derived from the partial order between object taxa. An action taxon *M* is a subtaxon of *N* if the description of *N* contains a subset of the taxa of *M* and every object taxon occurring in *N* is a subtaxon of a taxon in *M*. For example, a generalized action for a person to cut food with any knife and not restricted to any particular time and location is:

cutBread_2 :: {PhysObj+, Human+} x {PhysObj+, Tool+, Cutlery+} x {PhysObj+,
Liquid-, Tool-, Edible+} x {} x {} -> WorldState -> WorldState

This *cutBread_2* is a generalization, a superclass, of *cutBread_1* (*cutBread_2 > cutBread_1*) because the action taxon inheres the same action parameters and for each parameter, the taxon in *cutBread_2* is a supertaxon of the corresponding object taxon in *cutBread_1*: *Human* is a supertaxon to *Women*, because *{PhysObj+, Human+}* is a subset of *{PhysObj+, Human+, Female+};* the same taxon *Knife* is involved in both actions (note that "*a is supertaxon of b*" includes the case *a = b*). Finally, *Food* is a supertaxon of *Bread*, because *Food = {PhysObj+, Liquid-, Tool-, Edible+}* and *Bread = {PhysObj+, Liquid-, Tool-, Edible+, Dairy-}.*

In general terms, an action taxon *a* is a supertaxon of an action taxon *b*, if (after some reordering to establish correspondence):

$$a = o_1 \times o_2 \times \ldots \times o_n \to ws \to ws$$

$$b = p_1 \times p_2 \times \ldots\ p_m \to ws \to ws$$

$$a > b \quad \text{iff}$$

$$\text{for all } i < n: \quad o_i > p_i$$

$$\text{and} \quad n < m$$

7.3 Sensory-motoric program is abstraction from location and time

The generalization of an action to make it independent of location and time is a very common and important generalization of actions. A generalization of an action from location and time corresponds to a sensory-motoric program, which a person knows how to execute—we all have learned how to cut bread and are ready to execute this bodily knowledge any time any place where a knife and a bread is found.

This particular generalization is reflected in most languages I know of: verbs describe an action program, independent of location and time. "cutting" and "cutting bread" describe intensional sets of actions, which are action taxa in the above formalized sense, where "cutting" is a supertaxon to "cutting bread" and both are supertaxa to "cutting bread this morning", "I cut bread this morning in my kitchen", etc.

In the taxonomic lattice for actions, there is a branch for actions not localized in space and time and branches for localization and time. The combination of a generalized action with a location and a time is in the calculus of the taxonomic lattice a *meet* operation. The following Figure 5 shows the relevant parts of a taxonomic lattice for actions with the distinctions used so far; on the left side, the sublattice of actions by humans (independent of time and location) is shown, where we find *CuttingFood* and *CuttingBread* as two taxa; on the right hand side, the taxa that classify all actions *this morning* or all actions in *Geras*. The *meet* of these gives us the taxon of the actions *cutting bread this morning in Geras* (which may contain more than one execution and allow further distinctions).

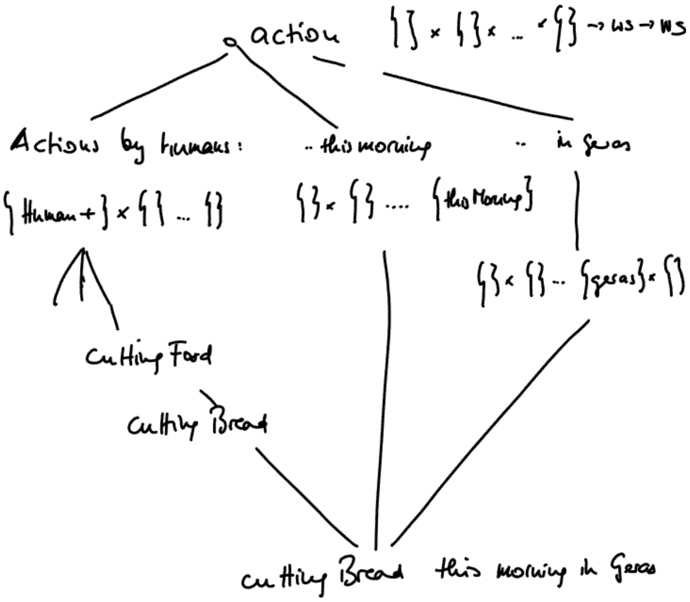

Figure 5: A section from a taxonomic lattice for actions

The common abstraction step for actions is to construct supertaxa by eliminating the time and the location where the action occurs; recall the rule 'reality is fully determined' that has us always start with a fully determined model, which is precise in space and time location. Many other generalizations are possible, for example the construction of the taxon for all actions executed by women:

done_by_female :: {Human+, Female+} x {} x {} … x {} -> ws -> ws.

The taxon for all *cuttingBread_by_women* would be the *meet* of *done_by_female* and *cutBread_2*.

We can see that specialization and generalization is possible in every argument of an action. Taxonomic lattices for actions become so complex such that only parts can be shown graphically. The calculus however is not affected and the computation of *meet* and *join* has a complexity that is linear in the product of the (small) number of action parameters times the number of distinctions used.

7.4 Structure of action parameters

Actions are described by a number of objects that inhere in the action, which is different for different actions. For the calculus with taxa, action parameters that are not meaningful in an action are replaced with the empty set of distinctions.

In practice, only a small set of circumstances influence an action and are conceptually included in its description given the limits of human cognition (Miller 1956). Western languages have established patterns of parameters for actions, which are often used and which I have used in the above examples: actor, object tool used, location, time, mode, etc. The particulars of what parameters are considered are not influencing the general theory, they affect only the examples and how we talk about them.

7.5 Taxonomic lattice of actions

The n distinction and the m action parameter structure together determine the taxonomic lattice of actions. It is the product of m times the taxonomic lattice constructed from n distinctions. The object lattice is thus a special case of an action lattice for an action with a single parameter; perhaps, this is an echo to the sentence "This is a *LoafOfBread*". Similarly, relations were classified by a product of two taxonomic lattices for objects (which reflects sentences like "The book is on the table"). An action like "Mary gives Peter a piece of bread (now, here)" gives rise to a product of three times the lattice of the objects (ignoring time and place).

Fortunately it is not required to represent each element of the taxonomic lattice of actions, as little as all the integers are represented to compute with them: they must only be representable, which is the case for the taxonomic lattice of actions as well. For a computer representation of taxa for actions, a representation must be selected that assigns different representations (bit patterns) to every taxon we want to distinguish. This needs 2 bits per distinction for each object lattice ($2 * n$ bits) and for the action lattice, with m parameters $m * (2 * n)$ bits. Even with a sizable set of distinctions and a reasonable number of action parameters, the number of bits required to represent each taxa remains small.

7.6 Rules restricting actions

Rules restrict what actions are possible, similarly to the rules for object taxa. For example, one might restrict that all actors must have the distinction *Human+* (or a distinction *Agent+*), that move operations must have the distinction *Location+* for their target location, a "tool" action parameter used must have *Human-*, etc. All combinations that do not respect these rules are mapped to the action taxon *bottom* (contradiction) that contains no possible action.

7.7 Semantics of operations

A description of an action *moveInContainer* with an action taxa as a product of the object taxa gives

moveInContainer :: {} x .. {PhysObj+} x {PhysObj+, Container+} x …

This cannot be separated from *moveOutOfContainer*, which requires the same distinctions for the objects involved. The image schema (Johnson 1990) connects the two actions, and does not provide for distinctions between them. The difference between *moving into* and *moving out of* is shown only in an axiomatic description of the image schema by operations (Kuhn and Frank 1991). For example, the effect of moving an object o into a container c is a state of the world in which c is in o (i.e., *in (o, c)*) (Axioms *A1*).

$ws_1 = movIn\ o\ c\ ws_0$

$in\ (c, o, ws_0) = false$
$in\ (c, o, ws_1) = true$ (A1)

The distinct action *moveOutOfContainer* follows the axioms *A2*:

$ws_1 = moveOut\ \ o\ c\ ws_0$
$in\ (c, o, ws_0) = true$
$in\ (c, o, ws_1) = false$ (A2)

The distinction between the action *moveIntoContainer* and *moveOutOfContainer* is in their use of axiom schema *A1* or *A2*; this is added as a parameter to the taxa and gives then two abstract taxa, which capture the difference between moving in and moving out; all specific actions are sub-taxa of these two:

moveIntoContainer :: … {PhysObj+} {PhysObj+, Container+} {A1+, A2- }
moveOutOfContainer :: .. {PhysObj+} {PhysObj+, Container+} {A1-, A2 +}

(It might be possible to encode *A2+* as *A1-*, or to add dependence rules *A1+* -> *A2-*, *A2+* -> *A1-*, etc.).

Note that the taxa related to these actions and the taxa describing the relation *in* must correspond in the distinctions of the objects involved. This established a relation between image schemata that are typically related to relation taxa, the axiom of algebraic definitions of semantics and the action taxonomy. (This results in dependencies like (*A1+* -> *{Container+}*)).

7.8 Corresponding actions

This taxonomy of actions separates actions like "Mary gives Peter a piece of bread (here, now)" from "Peter gets a piece of bread from Mary (here, now)", which can be rephrased as "Peter receives a piece of bread from Mary (here, now)". At the level of a verbal account, the two phrases may be roughly synonymous, stressing different actions. At the conceptual level and the level of motoric programs, the action of giving is very different from the action of receiving (Picture 6). Verbal accounts assume that the hearer completes the account with his commonsense knowledge—when there is a giver then there must be some receiver. The taxonomy of concepts must differentiate these, when verbal expressions do not (see later section 8).

Picture 6: The actions of giving and receiving

7.9 Uniqueness of taxa by adding distinctions

The taxonomy for actions assumes that the distinctions applied to the taxa of individuals are sufficient to separate all actions. This is achievable, because additional parameters for an action introduce new distinctions. For example, adding a parameter for the mode of an action, where the distinctions are fast/slow, forceful/weak, etc. separates "to walk" from "to move", and finer distinctions are imaginable to further separate skip, hop, and other types of movements.

The general rule of taxonomy is: If two concepts should be differentiated then there must be a distinction that separates between the two.

8 Communication

Communication between humans is exchange of symbols that stand for mental concepts but not necessarily in a *1-1* relation. In this section I will only point to the major differences between mental conceptualization and a verbal expression and suggest some similarities between human verbal communication and the exchange of geographic data between organizations.

A human constructs an internal conceptualization, often called a mental model, of a situation, which is in terms of (very fine) taxa. The conceptualization serves, as was stressed by Pinker and Prince (2002, 246), to allow inferring many other properties that are common for entities in a certain taxon from a few observations that are sufficient to classify something into that taxon. The mental model is less detailed than reality and contains the elements deemed important.

A human being expresses (i.e., represents externally) an *account* of an action. An account is a representation of the mental concepts in an external form to help others to construct mental models that are in some respect and for some aspects similar to the mental model of the sender (imply that the sender has honest and veracious intentions). Grice has described these expectations as the implicatures of communication (Grice 1989). The account uses the names of taxa and the sender assumes that distinctions she associates with the name of a taxon corresponds (with respect to the intention of the communication) with the taxon the recipient associates with this distinction. Real communication often shows that this is not the case. Example: Peter asks for a chair. Dan shows him the object in Picture 5 (last). Dan says then something to the effect of "No, I need a chair, not a stool"—pointing out that he makes a distinction between a taxon *Chair = {… fourLegged+, Support+, Back+…}* and a *Stool = {… fourLegged+, Support+, Back-…}*, whereupon Peter brings him the object in Picture 5 (first) and probably adds the word 'stool' to his vocabulary. The exchange may be more detailed and points out the distinction important for one of the participants and which is added without mental effort to the taxonomy of the other party.

Communication between organizations follows similar patterns, but the amounts of data exchanged is typically larger and feedback therefore not feasible in detail. Integrating the distinctions used by each organization and the connection between labels for taxa and the set of distinctions clarifies semantics bottom up and with a large multiplier effect.

9 What Has Been Achieved?

9.1 Not linguistic but cognitive

The mental conceptualization and the semantics of words are optimized for different goals. A taxonomic lattice allows mental inferences; verbal accounts communicate with minimal effort as precisely as necessary. Most of ontological and taxonomic work is oriented towards words, and expressions from natural languages. Expressions in natural languages are practically important and easy to observe. It seems important to identify the methods used to construct mental concepts and to separate them from the semantics of words. Distinguishing verbal expressions and concepts might also be useful to understand better the prototype effects in radial categories (Rosch 1973)

9.2 Extensible set of concepts

If in a situation a new distinction becomes necessary, the corresponding set of concepts emerges automatically. Adding a single distinction to a taxonomic lattice of objects multiplies the number of distinguished taxa by a factor of 3, but many may be ruled out by dependencies between concepts. This is an economic and powerful method to build concepts with detailed semantics. Language has numerous constructions to follow and to further divide the semantics of a word by composition.

The development of science and technology makes it regularly necessary to add new distinctions. Example: when cars where invented around 1886, the distinction 'without horses' became meaningful for carriages and concepts for motor cars as 'carriages without horses' were immediately available. Adding a distinction "with draftAnimal/horseless" for carriage produces immediately concepts like 'Garage' as place to store (horseless) carriages, and driving as an action concept ready to be applied to horseless carriages as well.

Application of a distinction to a concept where this distinction was not customary before is possible without explicit addition of concepts: when the distinction 'red' was used in conjunction with the taxon 'meat' and the new taxa 'red meat' and 'non-red meat' emerged, they were immediately understandable as concepts without an operation 'add "red meat" to list of concepts' performed by every American.

9.3 Integration of axiomatic definitions of actions with experiential distinctions

The calculus described here integrates in a single formalism the definition of semantics using algebraic approaches with experiential distinctions sometimes referred to as Image Schema (Lakoff 1987) and closely related to affordances. Gibson has introduced affordances as a shorthand description for what in a situation can be done with an object: a doorway affords to walk through it, a handle affords to turn (Gibson 1986) and Douglas Norman has further explored how well design communicates the action the object affords (Norman 1988).

An action taxon can be associated with an axiomatic schema (e.g., the container schema); subtaxa for a taxon that contains only this axiom all describe operations that have the semantics defined by the axiom. These axioms can be used at a high level of abstraction to plan which type of action to select and how to combine them. Such a taxon that defines semantics only by an axiom is abstract and contrasts with the less abstract taxa that represent a sensory-motoric program.

The axioms describe relations that hold before and after an operation is executed (compare the pre- and post condition of program verification in the Hoare style (Floyd 1985; Hoare, Hayes et al. 1987). Axioms refer to relations that hold between objects, for which distinctions identify operations possible. E.g., the distinction *Container+* refers to the execution of an action *moveInContainer*. The physical preconditions for an action—which is essentially what Gibson’s affordances are—can be seen as distinctions; a certain shape and mechanical construction produces things that can be pushed or turned—two different affordances, but also two distinctions. These distinctions are manifest in the form and observable—Norman’s "knowledge in the world" (Norman 1988). The point is the connection between the form, mechanical set up (and other functional structural elements) and the actions one can perform with an object. The distinctions connect between relevant object properties—which are part of a static world—with the dynamic world of actions.

9.4 Taxonomies ordered by level of detail

The level of detail of a taxonomic lattice is given the set of distinction it is constructed from. This gives a partial order between taxonomic lattices—the order relation is the subset relation between the set of distinctions they are constructed from. A taxonomic lattice *A* is a less detailed one than a taxonomic lattice *B* if the set of distinctions from *A* is a subset of the set of distinctions of *B*.

The set of distinctions produces all concepts included—they are of a single level of detail, even if not everywhere the finest distinctions are necessary. Adding a distinction if necessary is monotonic—no previous concepts or arguments becomes invalidated. Multiple level of detail can coexist in a form that for certain subsets not all levels are used (but they would be there if necessary).

9.5 *Join* and *Meet* always defined

The *meet* is the answer to the question what is the least specific action or object that has all the distinctions of the given ones. It gives the answer to a question like what has the form of a sexy young blond lady {*Sexy+*, *Blond+*, *HumanForm+*} and is produced by industry {*MassProduction+*}; the result is *Barbie_Doll*. It can also be used to find an action as the *meet* of '*move x from A to B*' and *A* and *B* having distinctions {*waterLocation+*} then the answer maybe *swim*.

Acknowledgement

This paper has emerged from a protracted effort to understand types in ontologies, databases, and program languages. Werner Kuhn has often advanced my understanding and repeatedly asked important questions. I appreciate advice form my daughter Stella Frank to improve the presentation.

References

Belnap, N. D. (1977). A Useful Four-Valued Logic. Modern uses of multiple-valued logic. G. Epstein and J. M. Dunn. Dordrecht, NL, Reidel**:** 5--37.

Claramunt, C. and M. Thériault (1996). Toward semantics for modelling spatio-temporal processes within GIS. 7th Internation Symposium on Spatial Data Handling, Delft, (August 12-16, 1996), IGU.

Fellbaum, C., Ed. (1998). WordNet: An Electronic Lexical Database. Language, Speech, and Communication. Cambridge, Mass., The MIT Press.

Floyd, C. (1985). On the Relevance of Formal Methods to Software Development. Formal Methods and Software Development. C. F. H. Ehrig, M. Nivat, J. Thatcher, Springer-Verlag. **2:** 1-11.

Fodor, J. A. (1984). "Precis of The Modularity of Mind." Behavioral and Brain Sciences **8**: 1-5.

Fodor, J. A. (1995). The Elm and the Expert: Mentalese and Its Semantics, MIT Press.

Fonseca, F. T., M. J. Egenhofer, et al. (2002). "Using Ontologies for Integrated Geographic Information Systems." Transactions in GIS **6**(3): 231-57.

Frank, A. (to appear 2006). Distinctions Produce a Taxonomic Lattice: Are These the Units of Mentalese? International Conference on Formal Ontology in Information Systems, Baltimore, Maryland.

Gangemi, A., N. Guarino, et al. (2002). Sweetening Ontologies with DOLCE. EKAW 2002.

Gärdenfors, P. (2000). Conceptual Spaces, MIT Press.

Gibson, J. J. (1986). The Ecological Approach to Visual Perception. Hillsdale, NJ, Lawrence Erlbaum.
A classic. Gibson introduces the concept of "affordances" that is central to interface metaphor theory, as well as a sort of a "user centered" view of space and its elements.

Grenon, P., B. Smith, et al. (2004). Biodynamic Ontology: Applying BFO in the Biomedical Domain. Ontologies in Medicine. D. M. Pisanelli. Amsterdam, IOS Press**:** 20-38.

Grice, P. (1989). Logic and Conversation. Studies in the Way of Words, Harvard University Press**:** 22-40.

Gruber, T. (2005). "TagOntology - a way to agree on the semantics of tagging data." Retrieved October 29, 2005., from http://tomgruber.org/writing/tagontology-tagcapm-talk.pdf.

Guarino, N. and C. Welty (2000). Ontological Analysis of Taxonomic Relationships. Proceedings of ER-2000, 19th Int. Conference on Conceptual Modeling. E. Laender and V. Storey, Springer-Verlag.

Hoare, C. A. R., I. J. Hayes, et al. (1987). Laws of Programming. ACM Communications.

Hornsby, K. and M. J. Egenhofer (1997). Qualitative Representation of Change. Spatial Information Theory - A Theoretical Basis for GIS (International Conference COSIT'97). S. C. Hirtle and A. U. Frank. Berlin-Heidelberg, Springer-Verlag. **1329:** 15-33.

Jackendoff, R. (1996). Languages of the Mind. Cambridge, MA, MIT Press.

Jensen, K. and N. Wirth (1975). PASCAL User Manual and Report. Berlin-Heidelberg, Springer-Verlag.

Johnson, M. (1990). The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason, University Of Chicago Press.

Kuhn, W. (1994). Defining Semantics for Spatial Data Transfers. 6th International Symposium on Spatial Data Handling, Edinburgh, UK, IGU.

Kuhn, W. and A. U. Frank (1991). A Formalization of Metaphors and Image-Schemas in User Interfaces. Cognitive and Linguistic Aspects of Geographic Space. D. M. Mark and A. U. Frank. Dordrecht, The Netherlands, Kluwer Academic Publishers**:** 419-434.

La Palme Reyes, M., J. Macnamara, et al. (2002). Count Nouns, Mass Nouns, and Their Transformations: A Unified Category-Theoretic Semantics. Language, Logic, and Concepts. R. Jackendoff, P. Bloom and K. Wynn, MIT Press**:** 427-452.

Lakoff, G. (1987). Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. Chicago, IL, University of Chicago Press.

Masolo, C., S. Borgo, et al. (2003). WonderWeb Deliverable D18 (Ontology Library). Trento, Italy, Laboratory For Apüplied Ontology - ISTC-CNR**:** 247.

Miller, G. A. (1956). "The Magic Number Seven, Plus or Minus Two; Some Limits on Our Capacity for Processing Information." Psychological Review **63**: 81-97.

Norman, D. A. (1988). The Design of Everyday Things. New York, Doubleday.

Parnas, D. L. and J. E. Share (1978). Language Facilities for Supporting the Use of Data Abstraction in the Development of Software Systems. Washington, D.C., 20375, Information System Staff, Naval Research Laboratory.

Peterson, J., K. Hammond, et al. (1997). "The Haskell 1.4 Report." from http://www.haskell.org/report/index.html.

Peyton Jones, S., J. Hughes, et al. (1999). "Haskell 98: A Non-Strict, Purely Functional Language." from http://www.haskell.org/onlinereport/.

Pinker, S. and A. Prince (2002). The Nature of Human Concepts: Evidence from an Unusual Source. Language, Logic, and Concepts. R. Jackendoff, P. Bloom and K. Wynn, MIT Press**:** 221-261.

Raubal, M. (2000). Ontology and Epistemology for Agent-Based Wayfinding Simulation. EuroConference on Ontology and Epistemology for Spatial Data Standards, 22-27 September 2000, La Londe-les-Maures, France, Institute for Geoinformation, TU Vienna, Austria.

Rosch, E. (1973). "Natural Categories." Cognitive Psychology **4**: 328-350.

Rosch, E. (1973). On the Internal Structure of Perceptual and Semantic Categories. Cognitive Development and the Acquisition of Language. T. E. Moore. New York, Academic Press.