CERIAS Tech Report 2005-71

THE ONTOLOGY OF EMOTION

by Katrina Triezenberg

Center for Education and Research in
Information Assurance and Security,
Purdue University, West Lafayette, IN 47907-2086

The Ontology of Emotion

A Thesis

Submitted to the Faculty

of

Purdue University

by

Katrina Triezenberg

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

September 2005

CHAPTER 1: ONTOLOGICAL SEMANTICS

Ontological semantics, as envisioned by Nirenburg and Raskin (2004), is a system of software and databases that contains enough information to competently parse both the syntax and the semantics of natural language texts. The book discusses the fine points both of the system's philosophy and of its inner workings, and the interested reader should refer to it for further reference. What follows is a layman's introduction to the system and to its possible applications, limited to what it is necessary to understand for the purposes of this particular effort and avoiding any serious discussion of the several points of divergence, often through upgrading, between this introduction and the 2001 snapshot of ontological semantics captured in the book.

As stated above, the system consists of software and of databases. The latter are called the static resources, because they are used, as they are at the moment, every time a text is processed. The various pieces of software, most importantly the analyzer and the generator, are the dynamic resources because trivially they transform texts into text-meaning representations (TMRs) or the other way around. Nothing is truly static in ontological semantics. One assumes that the system will be constantly dealing with brand-new texts. Calling even the databases the "static" resources is also somewhat deceptive, because the databases themselves are constantly changing and being added to, and are therefore not truly static. At the current time and for a long time into the future,

the databases will be undergoing massive expansion by trained acquirers with the ultimate goal of amassing enough information to parse the meaning of anything that makes sense to a human; there is an amount of primary information that is urgently necessary for this to happen, and it is the goal of acquirers to discern and to organize it. Should the project ever move beyond this primary phase, it will continue to gather infinitely specialized information, both moving into obscure domains of knowledge and re-working the existing knowledge into a finer and finer grain size. No human being ever knows everything about the world, not only because he cannot be exposed to everything, but because the information in the world is constantly changing and expanding—so does the information in the ontological databases.

The components of the system are as follows. The software includes:

1) An ecological analyzer for each language the system knows, which strips an incoming text of punctuation and other print conventions, such as the marking of numbers *(No., #, etc*.);
2) A tokenizer for each language the system knows, which breaks the text into its component morphemes;
3) A text analyzer for each language the system knows, which picks out the verb/argument structure and all such syntactical pieces from the text and transforms the text into its TMR;
4) A text generator for each language the system knows, which is capable of doing all of this in reverse, thus building a text up from its TMR.

Most of these pieces of software are not particularly new or exciting to the linguist, who was basically ready to write them before there was such a thing as software

to write. What is exciting and unique about ontological semantics are the databases, which allow the system to analyze and generate not just the syntax of a text, but its semantics also. There are four kinds of databases:

1) An ontology, which is a catalogue of universal, language-independent concepts. Entries in the ontology represent ideas rather than labels, and the ideas span many cultures. Many, many peoples of the world have an idea of a knife, of the color red, of a pet cat, and of freefall—each people has its own separate word or words, but the general idea is the same. This is what the ontology catalogues.
2) A lexicon for each language the system knows. The lexicon is a dictionary of a language. Instead of using plain text definitions the way an ordinary dictionary does, though, a lexicon explains the meaning of a word by linking it to concepts in the ontology. The definition for "Siamese cat," for example, will link the word to the concept for "pet cat," and probably note that its body (a concept) is white (a concept) and its tail, paws, and ears (all concepts) are brown (a concept). How precisely this is done will be explained later, when we look at the anatomy of ontological and lexical items.
3) A fact database contains information about things that have actually happened in the world. For example, the ontology might contain the concept of "Olympic games," the English lexicon would link the word "Olympic games" to that concept, and the fact database would list what years and in what locations specific Olympic games were held, who competed in what events, and what their scores were.

4) An onomasticon is a dictionary of proper names. It is rather similar to the fact database, but its information is about who and what exists, rather than what has happened. The onomasticon might know that a person called “Happy Harada” existed; the fact database would know that he was an Olympic ski-jumper. The onomasticon would also know the names Lillehammer and Nagano, but the fact database would know that Olympic games were held in these places and that Happy Harada competed in them.

The reader who is familiar with object-oriented programming will recognize in these databases the idea of instantiation; that is, one part of the system provides a general framework for what kinds of things can exist (the ontology), and then the same framework is used to create multiple, unique instantiations of the same thing (the onomasticon). The thing, being instantiated, then begins to interact with its environment (the fact database). Put practically, the ontology provides a framework for what a CAT should be like. The onomasticon then creates instantiations of CAT called Felix, Garfield, Nermal, and Tom. What these four unique instantiations do is then recorded in the fact database—Felix’s likeness is turned into a pop-art clock, Garfield eats lasagna, Nermal is cute, and Tom chases Jerry (an instantiation of the concept MOUSE). Note that the lexicon serves only to translate texts from human-tractable natural language into the computer-tractable ontology.

The true dynamic resources, as we have said, are the individual texts that the ontological system processes. The input texts themselves are therefore dynamic resources; so are the output texts, created by the text generator software. There is a third dynamic resource, which is less intuitive but crucially important, and it is the Text

Meaning Representation, or TMR. The TMR is what is produced by the syntactic/semantic parser, and it is a computer-tractable representation of what precisely the text means, not in natural language but in a complex branching hierarchy of ontological concepts and temporal and aspectual markers. If two texts that mean exactly the same thing are put through the system, the system will produce the same TMR for both texts. Such variety is unfortunately not possible in reverse—if you feed those two identical TMRs into the text generator, it will generate the same text from them. Such are the practical limits of computation. It would be fun if the text generator could be expanded to produce texts in dialects or attitudes, but, for the time being, attention is focused on expanding the databases…

In the ontology, each concept has its own entry. The flat version of the ontology is therefore a very long list of entries. Each entry contains the following parts:

1) The title of the entry itself. This is a label, or handle, that was attached to the concept so that the concept can be handled. Because the ontology has been developed in the United States, the handles are all English words and phrases. In some cases, the correspondence between the concept and its handle is synonymous, for instance in the concepts labeled CAT and KNIFE. The English words “tiger”, “puma”, “bobcat”, “Siamese cat”, “tabby”, “puss”, “kitty”, and “puddy tat” would all also be linked to the concept CAT. Some concept labels are phrases, such as CHANGE-STATE-OF-MATTER and ANIMAL-DISEASE. It is extremely important to understand the difference between the concept itself, which is language independent, and the label chosen for it.

2) An ontological concept will usually have a field that contains a natural language definition, in order to explain it to any humans who may be perusing the ontology. The definition means nothing to the computer and plays no role whatsoever in any computational process; an acquirer who spends time writing a definition field for an ontological entry has accomplished nothing. The definition field is, basically, a scratchpad.

3) The Is-a field lists the ontological concepts from which the concept is descended. The ontology is organized as a hierarchy with multiple inheritance, and a concept inherits all of the information about all of its ancestor concepts—the power and mystique that multiple inheritance lends to the ontology will be discussed later, but for now, understand that the Is-a field of the concept CAT might contain MAMMAL as well as LARGE-ANIMAL, SACRED-ANIMAL, and any number of other appropriate concepts. The only ontological concept that does not have an Is-a field is the root node ALL.

4) The Subclasses field, which lists all of the child concepts. An ontological concept might have no children, in which case it is called a terminal node. If it has any children at all, it must have at least two, never just one. If there is only one child concept, then it isn't really a child concept at all because it stands in contrast to nothing, and one is better off just using the parent. The concept CAT might have in its Subclasses field WILD-CAT, LARGE-CAT, PET-CAT, ASIAN-CAT, and any number of other more or

less relevant and detailed categories, depending on the grainsize to which is has been developed.

5) The Lexicon field, which lists all of the lexical items which have been linked to the ontological concept. In the entry for CAT, the lexical field would contain all of the list of tiger, puma, bobcat etc. from above as well as many, many others—this assuming of course that a more specific concept has not been created; if the concept for CAT has a child concept WILD-CAT, then WILD-CAT would be a more appropriate node for tiger, puma, and bobcat to be linked to.

6) The rest of an ontological entry consists of property slots. Slots contain the semantic information about a concept.

When we say that a concept inherits information from its ancestor concepts, we mean that it inherits their slots. In order to justify the existence of a concept, though, it needs to contain some information that distinguishes it from its ancestors—it needs either to have new slots, or to not inherit some information from its ancestors.

For example, the concept MAMMAL will inherit from its ancestor VERTEBRATE the information that a mammal must have a backbone made of true bone. It is then distinguished from its ancestors and its ancestors children (other children of VERTEBRATE would be FISH, AMPHIBIAN, REPTILE, and BIRD) by the fact that members of MAMMAL have either hair or fur, give birth to live young via true placental gestation, and produce milk. This is an example of a concept being distinguished from its ancestors via new slots; because nothing else in the world involves hair or fur or placental

gestation, slots containing information about these things must be created fresh for MAMMAL.

Underneath the concept MAMMAL, however, the dedicated biologist must eventually create a category for the MONOTREME, that troublesome category of mammal which lays eggs, and therefore does not bear young via true placental gestation. Oops indeed. Monotremes have been determined to be mammals, but some information which applies to all other mammals does not apply to monotremes—and so the concept MONOTREME will be stripped of any information about placental gestation. This is an example of a concept being differentiated from its ancestors by removing information.

One understands that the acquirer could, instead of putting the information about placental gestation into the entry for MAMMAL, put the information into every child concept of MAMMAL except the concept MONOTREME. This would be tremendously more work, however; monotremes are an extreme and specific exception to a characteristic that is, otherwise, definitive for the group. It is therefore most appropriate in this instance to include information in the ancestor, and remove the information from the one child to which it does not apply (but to which all of the other defining characteristics—hair, fur, milk, do).

A slot has three parts. First, the slot itself; second, the facet; third, the filler.

The slot contains one of the descendants of the concept PROPERTY. Properties are quasi-concepts that are important in describing other things, but do not stand by themselves. Beauty, age, goodness, size, and order are all properties, and any number of things also. So, a slot declares that the concept it is modifying is capable of having a certain property. Properties can be divided into two major groups, the relations and the

attributes. If the property is a relation, then it relates two ontological concepts via the slot; that is, it relates the concept in which it is a slot to the concept that is the filler in its slot. The property CONNECTED-TO is a relation; when it is a slot in the concept MOUTH, its filler must be ESOPHAGUS. It relates the two concepts to each other. One very important branch of RELATION is CASE-ROLE. The case roles are equivalent to theta roles in syntactic theta theory, and are used to structure verb arguments in the TMR. The case roles available to the system as this is being written are SOURCE, PATH, INSTRUMENT, DESTINATION, AGENT, ACCOMPANIER, BENEFICIARY, EXPERIENCER, LOCATION, PURPOSE, and THEME. There is also a complete set of inverse case roles.

An attribute, in contrast to a relation, comes bundled with a fixed number of possible fillers that are not concepts. Some attributes are binary, such as gender MALE/FEMALE, and some attributes have several possible fillers, such as COLOR-ATTRIBUTE, which has as its set of possible fillers WHITE TAN PURPLE ORANGE NAVY-BLUE GREEN CYAN BROWN BLACK BEIGE BLUE COLORLESS GRAY MAGENTA OFF-WHITE PINK RED VIOLET YELLOW. The story of how this particular set came to be is an interesting one with significant ramifications for the acquirer, and is told in chapter 2.

The facet describes in what way the filler applies to the slot. There are a small number of possible facets, and they are components of the software; that is, they cannot be added to or subtracted from or changed in any way by the acquirers, unlike both ontological concepts and attribute fillers. The complete list of facets, and explanations of

what they mean, may be found in Nirenburg and Raskin (2004: 198-201). The most commonly used facets are VALUE, SEM, NOT, and INV.

The structure of a lexical item differs from the structure of an ontological item primarily in lacking the inheritance slots Is-a and Subclasses, and in having multiple senses within a single entry. Its principal parts are:

1) The title of the entry, which is a graph. Underneath this entry is included the complete set of homographs—all of the different meanings. The entry for "bear" will therefore include the noun meanings "a large furry predator," "a cuddly child's toy," and "a grouch," as well as the verb meanings "to carry," "to tolerate," and "to give birth."
2) A title for each of the senses, followed by a marker that indicates the lexical category of the sense plus a unique number. The three noun senses of bear would be marked bear-N1, bear-N2, and bear-N3. The three verb senses would be marked bear-V1, bear-V2, and bear-V3.

These two parts alone may be said to constitute a lexical entry, which is a skeleton structure. The real meat and potatoes are in the senses themselves. The structure of a sense consists of four parts: a title, an ANNO, a SYN-STRUC, and a SEM-STRUC.

1) The title. This line assigns a lexical category and a number to the sense. This line also includes a redundant field that notes the lexical category (CAT N). It also contains a field (SYNONYMS "")in which all of the synonyms can be listed.
2) The ANNO: This part may be excluded from any sense, if there is no use for it. Any of the individual fields may be left out, also. All of the fields in the ANNO are there entirely for the convenience of the lexical acquirer, and play no

computational role at all. The field (DEF “”) is used to write a plaintext definition of the sense. The field (COMMENTS “”) is used to jot down any notes the lexical acquirer has, and is commonly used to note whether a verb sense is transitive or intransitive. The field (EX “”) is used to show examples of the word being used in a phrase.

3) The SYN-STRUC: This part tells the computer how the word is used in a sentence. In the case of nouns, there are no special instructions for usage, so the SYN-STRUC merely notes that the sense is used as a noun. In the case of verbs, information about what kinds of arguments the verb can take is noted here. In the case of adjectives, what kinds of words it modifies is noted here.
4) The SEM-STRUC: This part contains the semantic information about the sense. In one case, the sense lead-N2 maps directly onto the ontological concept LEAD, so no further information is necessary. Rather few senses map directly onto an ontological concept, however, so the SEM-STRUC usually contains a number of slots, which are just like ontology slots. They contain the slot itself, which is a PROPERTY from the ontology, a facet, and a filler.

An ontological concept may look like this:

(animal-mineral
 (definition (value (a mineral used as a nutrient by animals)))
 (is-a (value (animal-substance element-material)))
 (subclasses (value (iron zinc potassium calcium sodium phosphorus copper molybdenum iodine fluorine chromium selenium manganese)))

(theme-of (value (ingest)))

(inside-of (value (food animal-part)))

)

And a lexical item may look like this:

```
(MIND
(MIND-V1 (POS V) (ANNO (DEF "You mind your mother!") (EX "") (COMMENTS
       ""))
       (SYN-STRUC
       ((ROOT $VAR0) (CAT V) (SUBJECT ((ROOT $VAR2) (NP CASE
       NOMINATIVE)))(DIRECTOBJECT ((ROOT $VAR3) (NP CASE ACCUSATIVE)))))
       (SEM-STRUC (OBEY (AGENT (VALUE $VAR2)) (THEME (VALUE $VAR3) (NOT
       EVENT)))))
(MIND-V2 (POS V)
       (ANNO (DEF "to be bothered") (EX "He minds my whistling all the
       time")(COMMENTS ""))
       (SYN-STRUC
       ((ROOT $VAR0) (CAT V) (SUBJECT ((ROOT $VAR2) (NP CASE
       NOMINATIVE)))(DIRECTOBJECT ((ROOT $VAR3) (NP CASE ACCUSATIVE)))))
       (SEM-STRUC DISTURB (THEME (VALUE ^$VAR2)) (BENEFICIARY (VALUE
       ^$VAR3)))))
(MIND-N1 (CAT N)
       (ANNO (DEF "intelligence") (EX "Use your mind and you'll figure
       it out!")
       (COMMENTS "I'm not mapping to BRAIN because we're not talking
```

```
about the mass of tissue and nerves when saying 'mind',
usually"))
(SYN-STRUC ((ROOT $VAR0) (CAT N)))
(SEM-STRUC (INTELLIGENCE))))
```

These two examples are taken directly from the actual ontology and from the English lexicon. In them, one begins to see some of the problems and issues of ontological and lexical acquisition, which will be delineated in chapter 2.

The most obvious and straightforward application of the ontological semantic system is semantically-correct translation between two languages. One feeds the system a text in Language A, it is filtered through the ecological analyzer, tokenizer, parser, and then into a TMR via the ontology—that is, via a deep semantic understanding of the words and phrases in the text and how they relate to each other. From this TMR, one can then rebuild the text in the language of any lexicon available to the system, and the resulting text in Language B will be not a word-for-word translation of the text in Language A, but a meaning-for-meaning translation. It is important to understand that the lexicons are not just lists of single words, but also of compositional phrases. The system will not be left to puzzle out the precise meaning of “dog house” from its component words “dog” and “house”—the system would have no way to determine if the phrase means a house shaped like a dog, a house for dogs, or a house made out of dogs, nor would it have any information about the standard expectations of what a doghouse is like, for example that it is quite small, that it has no floor, and that it sits somewhere in the back yard. Likewise, the system will not be left to figure out what the idiom “in the doghouse” means. If the system must puzzle out a phrase from its component parts, it

will attempt to do so; as a general rule though, phrases that are not strictly compositional are included in the lexicon.

There are many other applications of the system which involve manipulation of the TMR. Each time one feeds a particular text into the system, it will produce precisely the same TMR (assuming, of course, that the static resources have not changed—particularly that the lexicon and ontology have not changed). If one fiddles with the TMR, re-generates the text, and then feeds that generated text back into the system, one will get the altered TMR back. This makes texts which have been translated into TMRs tamper-proof; even the smallest change semantic change in a text can be computationally (that is, quickly and effortlessly) detected. On the flip side of this, one can alter insignificant parts of the TMR in recognizable and statistically improbable ways, and therefore semantically watermark the text—creating a watermark that is resistant to copying and reproducing, unlike traditional paper watermarks and any number of other schemes, such as changing the spacing between letters. Tamperproofing and watermarking are different sides of the same coin because in the first case, one wants to create an extremely fragile pattern in the TMR, and in the second case, one wants it to be as robust as possible (see Atallah *et al*.,2001, 2002 for further discussion).

The very creation of a lexicon attached to the ontology is an application in itself, because it requires that a semantically precise representation of the item be created. For many words and phrases this is not a trivial task; one sees that the senses of "mind" given in the sample lexical entry above could be very much improved, given enough time and effort (and perhaps a graduate seminar in Cartesian philosophy). The core lexicon of a language is pretty solidly embedded in the minds of the language's native speakers, and

has no doubt been hashed and re-hashed by any number of dictionaries over the past couple hundred years. Getting good semantic representations of old words is therefore not really an issue, unless they are undergoing rapid semantic change. What are tricky are new words such as slang and jargon. For instance, there is a whole vocabulary used by computer programmers, hackers, and cyberpunks, but it is so specialized and so new (relatively—time marches on) that many of the terms can mean more than one thing, which can result in communication glitches between different groups of programmers, say, between McAfee's people and Symantec's people, when a brand new e-mail worm is sweeping the internet. In such a situation it is crucial that a large amount of very precise data be gathered in a very short period of time. Without standardized terminology, accidents will happen. It is therefore useful for both groups either to share a single lexicon or—and this is where ontological semantics comes in—for them to write their own lexicons semantically linked to a common ontology, which will do the semantic translation for them, and ensure that both sides understand what is being said (see Raskin *et a*l. 2002 for further discussion).

Because the ontology and lexicon contain so much information, the system begins to be able to make inferences and to "intelligently" assess queries. This makes it a valuable basis for what used to be called the expert systems—computer systems that are capable of answering questions posed by humans in natural language. It makes the systems capable of answering questions in a more intelligent way than the bag-of-words and keyword search approach currently does. Expert systems that did their job really well—and a large body of knowledge such as ontological semantics resources provide

would be necessary for the job to be done really well—could be used as internet and library search engines, tour guides, customer service, and technical support.

The above has been a brief introduction to ontological semantics as envisioned by Nirenburg and Raskin (2004) and improved and upgraded since. It is set apart from other Natural Language Processing (NLP) projects because it is invested in building resources. The overwhelming majority of NLP systems rely on syntax and statistics in an effort to avoid building resources. Ontological semantics is also set apart from other ontologies by the depth of its semantic information. The ontology entry given above was constructed at a relatively crude grain size and does not include all of the concept's inherited slots—the concept therefore consists of several times more information than shown above (see HUMOR in Chapter 4), and is absolutely unique within the ontology. Nirenburg and Raskin's ontological semantics is an extremely expensive one. The acquisition of resources requires a great deal of time to be put in by a large number of well-trained acquirers with more than a passing knowledge of linguistics, of semantics, and of the words and concepts that are being acquired. The hope is, of course, that all of this time and money will eventually pay off in creating a lasting and invaluable semantic resource that may be put to any number of uses.

In the next chapter, we will discuss the ins and outs of acquisition, and some of the principal issues involved in it. Chapter three will examine a specific very complex and underdeveloped domain, that of emotion, with the goal of figuring out how to acquire it. Chapter four will consist of the ontological acquisition of the domain. Chapter five will triage the results.

CHAPTER 2: METHODOLOGY

The ontological-semantic processing system is, in some ways, a system of irreducible complexity. Each of the static resources by themselves make sense as a database of facts or concepts or words, but without any one of the components, the performance of the system is severely crippled (in the case of the onomasticon and fact database) or stopped altogether (in the case of the ontology and the relevant lexicon).

Teaching ontological semantics is therefore problematic, because no one component really makes sense until one understands all of the components. Ontological semantics has been taught for nearly ten years, and the ideal progression of topics has not yet been found. Invariably, the audience is handicapped by an aspect of the system that they have not yet been introduced to: either they are unequipped to fully appreciate what they are being told, or they are completely distracted by questions. Throughout this chapter, teaching ontological semantics is seen as identical to a methodology of ontological semantic acquisition (see Triezenberg and Hempelmann 2005).

The standard progression of topics for teaching ontological semantics (and it is admittedly problematic) has been generally established as this:

1) Introduce the individual components
2) Introduce the path of data flow
3) Introduce the structure of concepts and of lexical items

4) Explain the ramifications of deep semantic linking and of multiple inheritance

5) Outline various applications, to illustrate the possibilities of the system

The first three points are necessary and basic, and must be explicated first, even though the more alert members of the audience will immediately begin to ask questions regarding point four. Points four and five are, indeed, the most interesting part of any ontological semantics workshop, and both the audience and the lecturer are invariably happy to spend as much time as is available discussing them—but quite often a short exploration of them renders a better understanding of the first three points absolutely necessary. It is therefore ideal for a workshop in ontological semantics to include two "passes" through the instructional materials: an introductory pass which allows the audience to generate questions about the system and its possibilities, followed by some workshopping time during which they attempt to acquire concepts or lexical items, or test an application. The second pass reiterates the first three points, and benefits the audience enormously as it answers a lot of questions that, during the first pass, they did not know to ask.

The teaching of ontological semantics has been, up to this point, conducted as workshops and seminars, and has faced the problems and limitations mentioned above. A manual for the teaching of ontological semantics—specifically for ontological semantic acquisition—has the principal advantage of being a hard copy of information that the audience may read over and over. The audience may double back and re-read things as often as they please, or may skip forward and try to find answers to questions.

The first chapter of this dissertation gave a reasonably thorough and, hopefully, clear explanation of the first three points, and touched on points four and five. The rest of

this chapter will be dedicated to answering questions regarding points four and five—a list of questions generated both by the author, as a person who has extensive experience in ontological semantics, and questions generated by the audiences of workshops and seminars. The problem of what order the questions should be answered in remains, and so the author hopes that the reader will hang in there, have patience, and trust that relevant issues will eventually be acknowledged. The principal issues to be discussed are as follows:

1) Multiculturalism and the ontology: are there really any universal concepts, do all languages have words for all concepts, and will there be fair representation of the worldviews of all cultures in the ontology
2) In practical terms, how are ontological and lexical items acquired: how does the acquirer work
3) In abstract terms, how are ontological and lexical items acquired: how does one decide what one is going to acquire, and how does one prepare to acquire it
4) So when do we start acquisition?

Multiculturalism and the Ontology

The ontology is touted as a "language independent" representation of the world—as a database of pure concepts as understood by the human mind. When told this, the usual initial reaction is to ask "what world? Whose understanding?" Anyone with a passing familiarity to more than one language knows that there is not a standard set of

concepts that map precisely one onto the other, and that there is therefore no such thing as a perfect translation, and most certainly such a thing as a disastrous one. This is why new translations of foreign works of literature are so often introduced—there is always a new scholar who claims to have a finer, more perfect understand of the author's original intent; or on the flip side, who wants the author's work to be completely reinterpreted, quite possibly in a way the author never intended. Here, for example, are several translations of a line from Edmund Rostand's French play Cyrano de Bergerac.

The original French is thus:

Je vous en prie, ayez pitié de mon fourreau:
Si vous continuez, il va rendre sa lame!

This, in this novice's best straightforward translation, says:

I pray you, have pity on my scabbard:
If you continue, it is going to reveal its blade!

When these two lines are fed into AltaVista's BabelFish translator, they come out thusly:

Please, have pity of my sleeve:
If you continue, it will return its blade!

In Brian Hooker's famous 1923 translation, which is often read as a book:

Pray you, be gentle with my scabbard here—
She'll put her tongue out at you presently!—

And in Anthony Burgess's 1971 translation, often used on stage:

Have some consideration for my scabbard, pray.

She loves my sword and wants my sword to stay

Inside her.

And finally, Lowell Bair's 1972 translation for Signet Books:

Please have pity on my sword: if you don't stop shouting you'll frighten it out of its scabbard.

These examples are extremely useful in illustrating, first, how hopeless is semantically impoverished machine translation. The BabelFish translation is in no way a finished product ready for any kind of publication, and is not meant to be. It simply provides a human with enough information to (usually… sometimes) piece together what the original text was saying, in a very general sense. Metaphor and idiom are completely lost. Brian Hooker and Lowell Bair's translations are reasonably faithful, though both take advantage of metaphor to enrich the translation, in ways that Edmund Rostand did not. Finally, Anthony Burgess's rendering of the lines is grammatically and semantically unfaithful, though witty and a little shocking as Cyrano meant himself to be.

Aside from good and bad translations, and the pitfalls of metaphor and idiom, there is quite often the plain problem that Language B doesn't have a word or phrasal that adequately represents a word or phrasal in Language A. An explanation of greater or shorter length may be substituted when translating, or the idea may be glossed over. When working between cultures it is even possible that Culture B doesn't have a concept that Culture A does. A favorite example is the genders in Swahili. A speaker of a

European language is accustomed to working with two or three genders which roughly correspond to biological realities—male, female, and neuter. Swahili, unfortunately for these European speakers, has fourteen genders and no representation of gender as a biological reality. One of the Swahili genders is a category for "things that spread," such as vines and pools of water. Does an English speaker have such a category? Probably he has not, until he has heard this explanation of the category—but once given this explanation, which is not particularly abstract, he can easily create the concept. Some concepts are less easily grasped from culture to culture and sometimes the natives of a culture claim that outsiders will never really understand a concept at all. However, it is probably a safe bet that perfect understanding of a concept and the ability to distinguish a concept from other concepts must not go hand in hand. Thus, when building the ontology, one faces the problem of whose concepts are created and when.

The answer to the question of who-what-when is that the ontology should represent as many concepts from as many cultures as possible, because this will make it a more universal NLP system, better able to competently process the texts that will be presented to it—in precisely the same way that a person who reads widely and omnivorously becomes better able to understand everything he reads. The ontology is in no way restricted in its representations of possible worlds: it is not meant to represent the American world, or the world of European thought, or the First World. It is simply meant to represent the world, and the wonder of multiple inheritance allows it to do this.

For a single acquirer with strong opinions about how a certain branch of the ontology should be structured, multiple inheritance is mostly useful for criss-crossing between domains; for instance, one builds the ANIMATE branch confidently and with no

hesitation, according to the standard evolutionary taxonomy, and is comfortable in one's mind. One is then able to create further meaningful categories within the branch but independent of evolutionary taxonomy, such as DOMESTIC-ANIMAL. Some acquirers may have different ideas about how the ANIMATE branch would properly be structured. Possibly the person has different ideas about which animals are domestic; possibly the person's culture has other categories such as SACRED-ANIMAL which have no particular relationship either to biology or to domestication. This acquirer may use multiple inheritance to impose his own worldview upon the ontology, thus enriching it and widening the range of texts it is competent to process and without disturbing the information that already exists—thus creating an ontology that simultaneously represents multiple points of view.

This successive layering of categories and worldviews upon the ontology could, though, become a problem. Suppose that so many different categories are imposed on the data that the ontology stops being a "tree" in any kind of sense, and becomes a "web." Maybe someone has an issue with the inclusion of such intuitively inanimate things as barnacles being included in the ANIMATE branch, and so turns things upside down by making a certain branch of the animal kingdom that excludes barnacles into a parent-concept of ANIMATE. This has formed a semantic loop, in which one can trace inheritance from ANIMATE to whatever the "truly" animate branch of the ontology was to ANIMATE and on and on and on, just like Pete and Repeat sitting on a fence. Even if semantic loops are not formed, there could easily be so many parents of every concept that the system—or indeed a person with so much knowledge in his head—is unable to

easily and reliably know which particular parent (which particular sense of the concept, that is) should be referred to at any point in a text.

It is possible that this problem would not occur, though, and the anthropologist more than anyone else is in a position to be worried or not worried about it occurring. As discussed in Shaul and Furbee (1998: Ch. 4), folk knowledge often seems to pattern universally from culture to culture. In the domain of animals, it has been found that a language which has a term for "mammal" will also have a basic term for "wugs": bugs, worms, small reptiles, and all the assorted small creepy-crawlies; from this it is deduced that all languages have terms for "fish," "bird," and "snake." This implies that the categorization is somehow natural and right to the human mind. Would the ontologist not better serve himself and the system, in the long run, by kowtowing to these "intuitive" categories? Should the ANIMATE branch not divide into FISH, BIRD, SNAKE, WUG, and MAMMAL? One philosophy for organizing the ontology that is supported by Nirenburg and Raskin is to make the ontology mirror the mind of the "common man." To this end, Nirenburg disapproves of making HUMAN a child-concept of APE, the two being so strictly set apart in folk knowledge. Biologists (one biologist, actually) who saw the ontology which set HUMAN apart from other branches of ANIMATE, however, wailed and gnashed teeth and set things to "right." The accepted order of evolution, in which humans are indeed one of the great apes, is a view held by the "common book-learner" and will be represented in the texts produced by and for that person. Evolutionists and creationists—whose viewpoints are crudely represented by the two possible states of the ontology—may have their debate elsewhere; inside the ontology, both viewpoints can and should be represented, so that the system is competent to

understand texts written from both (and the ratio of creationists to evolutionists versus the ratio of creationist *texts* to evolutionist *texts* would be very interesting numbers to know indeed). Folk-knowledge would, in this particular instance, simply impose an additional structure upon the ontology, rather than prove to be any sort of unifier.

The turnout might be very different in a domain that has less "scientific" chest beating about it; for example, color terminology. Certainly the wavelengths of different colors may be measured, but these wavelengths are then described using hexadecimal RGB notation, or something similar, rather than actually standing as concepts in the human mind. Again referring to Shaul and Furbee (1998), color studies have discovered that the single-lexeme, non-metaphorical, non-borrowed color terms in languages almost always follow a certain pattern of proliferation. If a language has two native color terms, they stand for black/dark/cool and white/light/warm. If a third term is present, it is red. If a fourth term is present, it is either yellow or green/blue. If six terms are present, they are black, white, red, yellow, green, and blue. Orange and purple get added next, and from there brown, gray, pink, and so on with whatever color terms languages may have. It is interesting that, at the very beginning, this progression mirrors what is believed about infant visual development. At first, there is only a distinction between light and dark. The first discernible "color" is believed to be red. Thus the past fad for black and white nurseries, which has now given way to a market for black and white and red toys (see http://www.geniusbabies.com/noname23.html).

The ontology as it stands has a quite minimal treatment of color, not even a branch of the ontology but rather a single property, which can take the literal values WHITE TAN PURPLE ORANGE NAVY-BLUE GREEN CYAN BROWN BLACK

BEIGE BLUE COLORLESS GRAY MAGENTA OFF-WHITE PINK RED VIOLET and YELLOW. This assortment of color terms appears to be nearly random to an English speaker; why BLUE and CYAN and NAVY-BLUE? Why WHITE and OFF-WHITE? Nevertheless, we see that black and white are both represented, and red orange yellow green blue and violet. So are brown, gray, and pink. Cyan is perhaps a nod to the Roy G. Biv theory of color, or to the ontology's founders' native Russian, which has a term for blue and a term for light-blue. Off-white, magenta, and navy-blue may either represent native color terms in the languages of other acquirers, or be spurious additions that were useful to a single person at a single point in time, the state of the ontology being what it is (that is, it contains a lot of questionable inclusions and exclusions).

This schema for color representation is, at first, disappointing. The schemer and organizer in us all would much rather have found four properties, for dark/light, redness, greenness, and blueness… or for the pigment-lover, yellowness, cyanness, and magentaness (aha! This is why cyan and magenta are there). Having this assortment of properties would allow unique definition of ranges for all of the colors in the lexicon, and frankly, at some point in time this will have to be addressed (how, in the current state of the ontology, does one define chartreuse? The best effort would be to fill in COLOR-ATTRIBUTE with both green and yellow, and this just isn't quite right).What the reader of anthropological linguists appreciates in it, though, is the fact that at some point, the ontology-builders decided to eschew optics and pigments and sliding scales of wavelengths, and to just include basic color terminology. In this aspect of the ontology, the common man's common mind has triumphed.

In both of these cases, the ultimate conclusion is that folk knowledge is an important aspect of human cognition (indeed, *is* human cognition) but that rather than being the principal aspect of the ontology's structure, it should be represented within the framework of a larger, more all-encompassing structure. That humans seem to naturally sort animals into birds, fish, snakes, wugs, and mammals is interesting and useful and highly salient, even to people who are trained to make further distinctions; the fact that further distinctions are made, however, trumps. Humans recognize these basic categories of animals because these categories are, for the most part, real: the major chordates and everything else, slimy amphibians and small reptiles being relegated to the category of everything else. The ontology must have a finer grain size, though, because finer grain sizes exist and are commonly recognized. In the case of color, it is excellent and mete that a list of the colors which frequently merit native terms are listed as the literal values of COLOR-ATTRIBUTE. More needs to be done, however, to account for all of the metaphorical, borrowed, and multiple-morpheme color terms which are equally present in all languages and cultures. Even small children are conditioned to this, with boxes of 32, 64, 96 and more crayons, many subtly different from each other and likely to have metaphorical names such as "jungle green" "kelly green" and "malachite."

What the ontologist discovers is that folk knowledge should be represented, and can easily be represented because it is a primitive reflection of the way things "really" are: and by this is meant not the way things really, really are, which is unknowable, but rather the full extent of what *seems to be*, to those who specialize in something and make a career out of recognizing and differentiating aspects of it (for of course, we can't do any better than that). Less intuitive, more specialized systems and structures need to be

available in the ontology, however, if the range of human experience is to be adequately and fairly represented. Nirenburg and Raskin's mythical "common man" is therefore, ultimately, a mistake. When time and manpower are running short and a large grainsize will do (and these were the conditions in the early days of the system, when the project was to scrap together some kind of core ontology), then folk knowledge should definitely be appealed to as a quick-and-dirty foundation of knowledge. Sooner or later, however, reality and variety must come creeping in, and if the text processing system is to function competently within texts that represent reality and variety then it must have some more sophisticated knowledge base installed.

The presence of a large and largely unintuitive structure does not preclude the existence of smaller specific ones. The problem of the ontological tree eventually turning into an ontological web, or mishmash, is however less worrisome once one has discovered folk knowledge. Given the intuitive division of animals that occurs in so many cultures, it is, for example, highly unlikely that someone will come along and want to acquire a category that includes birds, barnacles, and certain varieties of bloodsucking flatworms. Given that it is true that there are nonsense categories such as this, even when absolutely every opinion held by absolutely everybody is taken into account, one concludes that true ontological mishmash will never be made, because however many categories do emerge will be legitimate and meaningful, and therefore useful in the system. If the system is legitimate and meaningful, then however complicated and however webbed, it will make sense.

How the Acquirer Works

As time has passed, the way that the acquirer works has changed. From early times there was a server-based piece of software called the Knowledge Base Acquisition Editor, or KBAE, through which all ontological and lexical acquisition was conducted. Because KBAE sat on a server, all changes to the databases were available to all acquirers the instant they were made, which went a long way towards avoiding duplication and negation of effort. The KBAE software itself did three more highly useful things: first, it tracked the ramifications of changes. That is, if a concept was added, the acquirer was notified that it had been added as a child of all its parent concepts. If a concept was deleted, the acquirer was notified of all concepts and all lexical items that would be affected by the deletion—which becomes extremely important if one is deleting a concept to which any lexical items at all have been linked, as those items would otherwise be left with no ontological linking and therefore no semantic content. KBAE also enforced the "rules" of the system, ensuring that slots were not applied in an inappropriate way. For instance, only a small subset of the branch OBJECT is capable of having the case-role AGENT; at various times this has included humans, animals, organizations, and natural forces, and it has been a constant battle to decide whether any of these should be excluded or included. In any case, one cannot fill an AGENT slot with any other OBJECT, and especially not with an EVENT. If someone tried to make HAMMER the agent of HIT, KBAE would disallow the change (it should be the INSTRUMENT). Thirdly, KBAE would display a concept's entire inheritance, making it quick and easy to the acquirer to see what needed to be added or changed. Acquisition

was therefore, for some acquirers, an exercise in gracefully bowing to rigid authority, but quite a lot of responsibility for doing things correctly was lifted from the acquirer's shoulders.

While KBAE was in use, revisions to its software were constant and for the most part useful. KBAE was always slow, often taking 15-20 seconds to make a change, but it was usable, especially if an acquirer worked in multiple windows at one time. And then, something went terribly wrong. Something happened to KBAE—and no one knows precisely what, or precisely who was the agent (or the instrument)—to make it take well in excess of a minute to make a change. Programmers have since looked at its code and called it "fifty thousand lines of the worst code ever seen," but some of the principal problems with it are as follows:

1) The ontological entry does not include a list of all lexical items linked to that entry. Every time KBAE displays a concept, it must search all of the lexicon files (and there are three sizable ones, in English, Spanish, and Mandarin) for all references to that concept.
2) The ontological entry does not include any of the inherited slots or any information about inheritance past the parent/child level. In order to display a concept's full ancestry and inheritance, then, the whole ontology had to be searched through several times.
3) All of these searches must be repeated when a concept is being changed, so that the system can warn the acquirer about the ramifications of the change.

Of course all of these operations were always being done, long before the processing time skyrocketed to over a minute; what changed to slow it so radically is therefore not clear.

In any case, KBAE has been determined to no longer be a valid tool for acquisition, and researchers have had to find other ways to work.

The KBAE software sat on top of the ontology and lexical databases, which each consist of a single file. These "flat" files are the basis of acquisition at the present time. They are written in nested-parenthesis notation (as was seen in the sample items from chapter 1) which is reasonably human-tractable and, when written by humans, reasonably easy to check using parenthesis-matching software such as emacs. Before acquisition begins, the prospective acquirer must become absolutely comfortable with the format of entries and with the pattern of parentheses within an entry. An acquirer works in his favorite word processor, then saves the work as a plaintext file (it is important to cut the text and paste it into a plaintext file—if one tries to save as plaintext from a fancier word processor, the word processor is likely to leave a lot of hidden characters in the document) and eventually merges it with the master file, alphabetizing the entries..

This method of working is not entirely disadvantageous, interestingly. The acquirer does of course suffer without the protective shell of KBAE: if multiple acquirers are working, they can quite easily step on each other's toes by duplicating work and by creating non sequiturs. This problem is largely solved by strict division of labor within a group and by the maxim that one "never deletes" a concept—at this point acquisition is strictly that, acquisition, and never revisions that involve deleting information. The acquirers are also prone to create slots and fillers that they shouldn't, though it can be said that if a properly trained acquirer—one who knows that a hammer should be an instrument and not an agent—puts a wrong filler in a wrong slot, then the restrictions on that filler and slot are probably unnecessarily limiting and ought to be relaxed.

The principal, and glorious, advantage of "flat acquisition" is the speed with which it can be accomplished. Cutting and pasting an entry as a whole, and changing its title and a few of its filler values, is wonderfully efficient compared to navigating KBAE's labyrinth of fields and buttons and forms. Even if KBAE had been able to make changes instantly, it would still have been much, much slower than flat acquisition. The loss of KBAE, though sad, has therefore not been a complete disaster for acquisition work, and constructing a replacement for it has not been on the list of priorities at all. At this point in time, creating another piece of server-based acquisition software is remote on the to-do list, and will be put off indefinitely.

As fast as flat acquisition is, flat perusal is a dismally slow process. The ontology's flat file runs into the hundreds of pages, the English lexicon's flat file into the thousands. When KBAE was lost, it was crucially important that a piece of software beyond word processor search functions be created, so that the ontology at least could be navigated quickly and intuitively. Thus, the instant KBAE was lost, KBAE-Avoider (or just "avoider") was written. Avoider is a GUI for browsing the ontology file in a meaningful way—one starts with the ontology's root node ALL and opens or closes successive branches according to his wishes (this process of climbing up and down the ontology is called sliding). The user can also perform string searches for items. Note that, because the ontology's concepts are labeled with approximate handles rather than exactly descriptive words and phrases (often because English has no exact word or common phrase to describe the concept), searching for the title of a concept is quite often a losing game. One first tries searching for a lot of synonyms that the concept may be labeled by. An acquirer who is familiar with the ontology will have noticed, in browsing it, that there

are certain phrasal constructions quite commonly used as concept names, and this acquirer may have some luck in constructing similar phrases. If a concept cannot be found at all through string searches, one must return to the ontological tree and begin to slide. The structure of the ontology is somewhat controversial, and multiple versions have been made for different applications, but in general the root node ALL branches into OBJECTs and EVENTs, sometimes with a third branch of PROPERTIES (which are sometimes placed under one of the other branches). One is usually searching for an object or event, so one chooses the correct branch. All other branchings are nearly as intuitive as the first; OBJECT for example branches into SOCIAL-OBJECT, PHYSICAL-OBJECT, and MENTAL-OBJECT. Sliding is therefore a series of semantic decisions which require that the acquirer already have quite a clear idea of where the concept he is after should live. With any luck, the ontology provides no pitfalls or false avenues to mislead the slider, and he will be able to determine confidently that the concept he needs either does or does not exist.

Besides facilitating these two methods of searching the ontology (and imagine how difficult they would be, using the flat file) KBAE-Avoider displays all of the information which exists in the concept's entry. Note that this is different from what KBAE displays. The Avoider does not show any information about inherited slots at all. It is the acquirer's responsibility to look up the entries of all of the concept's ancestors. One is also unable to browse the concept's family history past its immediate parent in the Is-a slot—to look at a grandparent, one must open the parent, then open the grandparent. This is not a difficult or time-consuming procedure, but one does lose the place of the original entry while doing so. The third major drawback of the Avoider is that it is

completely separate from the lexicon, and one is therefore unable to explore the lexical linking of a concept.

The Avoider is now several years old. In recent months, spurred by the prospect of large scale acquisition by multiple acquirers working in multiple groups, new pieces of software have been produced which serve several useful functions:

1) A parenthesis matcher, which will look through either an ontological or a lexical entry and tell the acquirer which ones have problems (written by Dr. Christian Hempelmann)
2) A LexAvoider, which facilitates lexical searches and displays the lexical entry's information and allows the acquirer to create new, though structurally limited, entries (written by Courtney Falk, 2005a)
3) A Compilation of Acquisitions Tool, or CAT, which merges new files into the legacy databases (Falk 2005b)
4) A limited browsing and acquisition tool for both the ontology and lexicon, originally designed for a group of acquirers in Russia and therefore called the Russian Acquisition Tool or RAT (Falk 2005c).

Given so much nifty software, the acquirer has little choice but to get to work. When constructing an ontology concept, the first step is to fill in all of the standard fields: the title, definition, Is-a, and Subclasses. The next job is to find enough slots and fillers to uniquely and adequately describe the concept; unique because it must be completely different from all other concepts in the ontology (which, given inheritance, mostly means it must be different from all of its sibling concepts) and adequate because if one is going to do the job, one may as well do it right. In learning what needs to go into a concept, the

acquirer should not be misled by the quality of many entries in the legacy ontology: the quality is dismal, and they were executed by acquirers with no expertise and little training. For most concepts, the acquirer ought to be able to find two or three very clever slots and fillers, or half a dozen moderately clever slots and fillers. The first sets of slots one examines are the case roles (for events) or the inverse case roles (for objects): the acquirer wants to describe who commits what event; using what; when; how; where. Careful consideration of case roles will almost always result in a satisfactory ontological concept. After this, if one is working with an object, there is a large collection of spatial properties that may be useful, such as CONNECTED-TO, UNDER, ABOVE, INSIDE, and BETWEEN. If the concept is an event, there is a parallel set of properties describing the chain of causation. After this point, if the acquirer wants more slots (and he should) he must rely on his general familiarity with what is available, to find the right slots. It is therefore very desirable that an acquirer be intimately familiar with what the PROPERTIES branch has to offer, so that he doesn't miss things when acquiring, and doesn't waste a lot of time trying to find just the right slot. Indeed, an acquirer who spends any amount of time in serious work will become very familiar with the properties—rather more so than with objects or events, and not only because there are fewer properties.

The acquirer who is working without benefit of the KBAE will probably be creating "flat" entries in a text file. This acquirer may as well take advantage of this fact, and acquire a lot of items at once that all have approximately the same format, and create a template. Templates are one of the most important strategies for high-speed acquisition, and using them is called "rapid propagation." They are most often used for lexical

acquisition, given that so many groups of words are so often similar. For example, an acquirer working in the domain of botany can construct a template for “wildflower” and then very, very quickly work up a list of hundreds of items using it, changing only a few pieces of information for each item such as the title, flower color, and height. This is basic database building at its best and in such cases, it would be madness to construct each lexical entry from the ground up, fiddling individually with the SYN-STRUCs and the parentheses (and, incidentally, it would lead to many more formatting errors). The lexical acquirer is therefore strongly advised to focus on a domain, acquire large groups of highly similar words, and create a template to do so. Here is a sample of a template, created to rapidly propagate the names of human disease:

```
(
( (CAT N) (SYNONYMS “”)
(ANNO (DEF "") (COMMENTS "") (EX ""))
(SYN-STRUC ((ROOT $VAR0) (CAT N)))
(SEM-STRUC (
(caused-by (sem ))
(instrument-of (sem ))
(duration-typical (value ))
(has-symptom (sem ))
(treated-by (sem ))
(curability ())
(survivability ())
(treatability ())
(affect-quality-of-life (value ))
))))
```

When acquiring ontological entries in a word processor (sometimes called acquiring “by hand”) it is easy to learn the structure of the entry and the rhythm of parentheses within each slot and the entry as a whole. Doing parentheses manually for a lexical entry is quite a different task, because of the far more complicated structure of a lexical entry—multiple senses within the main entry, and each sense with its own ANNO, SYN-STRUC, and SEM-STRUC, and each of these with multiple fields all requiring a

confusing assortment of parentheses and quotation marks. To speed up the process and to cut down on formatting errors, I have constructed a text file called Lexicon Building Blocks (Appendix A) which contains each basic component of the lexical entry, so that they may be copied and spliced into each other conveniently.

Lexical acquisition is also complicated by the very structure of lexicons themselves—by what counts as a word by itself and should therefore be added, and what is only a form of a word and should be left to the tokenizer. An important part of the text processing system is the tokenizer, which breaks words into morphemes. A lexical acquirer should keep this piece of software in mind while acquiring, because during runtime the tokenizer will recognize and "break off" many grammatical morphemes, leaving the root behind. There is therefore no need to acquire many words that are formed with morphemes—one only needs to acquire the root.

What the tokenizer does do:

1) In general, the tokenizer can be relied upon to create all of the forms of a regular word's paradigm. A lexical acquirer would therefore acquire the first-person singular present form of a verb, such as "walk," and rely upon the system to generate its variants "walks, walking, walked". Likewise, one need only acquire the singular version of regular nouns ("bird" but not "birds, bird's, birds'") and the attributive forms of regular adjectives ("young" but not "younger" or "youngest").
2) The tokenizer may be relied upon to know irregular forms of words. For example, one may acquire the verb "run" and rely upon the system to generate its forms "runs, running, ran" even though "ran" is irregular. In fact, in these cases the

tokenizer works from a list of known irregular words and their forms, and it is therefore likely that an irregular word is already in the system, so check before you acquire it. The tokenizer will also recognize that the paradigm for “goose” includes “goose’s, geese, geese’s”. However, because sex and age are not part of the noun paradigm in English, the system will not associate “goose” and its paradigm with “gosling” and “gander” and their paradigms. The three words need to be acquired separately.

What the tokenizer doesn’t do:

1) As noted above, the tokenizer will not recognize forms of a word that are outside the word’s normal paradigm, such as the variants “goose”, “gander”, and “gosling”.
2) The tokenizer should not be relied upon to deal with derivational morphemes, such as the prefixes dis-, dys-, in-, im-, ir-, il-, and un-, all of which create antonyms of adjectives.
3) The tokenizer may not be relied upon to deal with morphemes that change the grammatical category of a word, such as the morpheme “-ly” that forms adverbs from adjectives, or “-ness,” which creates a noun out of an adjective.
4) The tokenizer should not be relied upon to break apart the two halves of a compound word, whether the meaning of the compound is compositional or not. “Raspberry” is not compositional (raspberries have nothing to do with rasps), and it should be acquired independently of “rasp” and “berry.” A “doghouse” is compositional (it is a house for a dog), but should be acquired independently of both “dog” and “house.”

In all cases of things that the tokenizer should not be relied upon to do, please understand that the tokenizer *can* do these things; it is simply preferable to have a lexical acquirer deal with them instead. A lexical acquirer will create a far more satisfactory semantic profile of a multimorphemic word than the tokenizer will. If, at runtime, the software comes upon a word that is not found in the lexicon, it will attempt to generate the meaning of the word based on its component parts—and it is equipped to handle derivational morphemes and compound words, but it will not handle them so well as a lexical acquirer would. In the example of "doghouse," the system would merely recognize that the word means a house that is in some way associated with dogs. It will not know that a doghouse is usually outside of the people-house, that it is very small, or that it usually doesn't have a floor or windows. For this reason, it is better to create lexical entries for compound words than to rely on the tokenizer to do so. Words created with derivational morphemes are usually also sufficiently different from their roots to warrant an effort by the lexical acquirer, rather than leaving the words to be dealt with a runtime, which would slow down the system and generate less satisfactory results.

One more consideration in manual acquisition is the issue of complex events: events which can be said to be composed of several simpler events. Of course, any event can be divided into smaller and smaller components until one comes down to not atoms but to charged particles, and whether they attract or repel each other. An ontology of minimal grain size would do just this—and also a standard college degree in animal physiology (as the author well knows and regrets), which means that it is not unreasonable to expect the ontological system to know more than a little bit about chemistry and physics. Practically, however, there are many cases in which treating an

event as complex is not helpful in any way shape or form. Take the example of walking, applying it to whatever species is your favorite. Walking is composed of a great number of very fine muscular movements, operating on hair-trigger feedback loops. To fully describe the act of taking a step one must describe what every muscle, bone, tendon, and nerve is doing and why it is doing in response to every other muscle, bone, tendon, and nerve. Yes walking can be described in this way—but why? Even scientists who study walking in order to build robots that walk have ultimately found it to be worse than useless—a fruitless complication of what seems to be the simplest thing there is. So in this way, there is very seldom a call to make walking into a complex event. Practically too, in time and cost, the deep level of expertise needed to acquire walking as a complex event is enormous. In short: walking is a common concept that the human mind is enormously prepared to treat as a simple event. Something like the Krebs Citric Acid Cycle is neither a common concept nor something that can be treated as a simple event—and so if anyone ever bothers to acquire all the processes of cellular metabolism, it will not be a simple event.

Complex events are created by putting the slot HAS-PARTS into an event concept. The acquirer then lists all of the events that are components of the complex event as fillers. Please note that for all slots except binary properties, the acquirer may list as many fillers as he pleases, and that these fillers may be ontology slots themselves—the slots may be nested. This means that inside of the HAS-PARTS slot, each event may be filled out with case roles; this can create quite a complete description of an event, including objects and people who take part in it.

When bothering to construct these complex event entries (sometimes called scripts), the acquirer should look carefully at the entries for all of the events that he intends to use as components. It is quite possible that they already have their case roles adequately filled, in which case he does not need to duplicate the effort, because all of that will be invoked automatically by the system. It is also possible that one of the component events he intends to use has been written as a complex event itself, and contains some of the other components he had intended to use. These issues are called "script hierarchy" issues, and they are rather a sticky business—in the grand scheme of things, the universe is one big script, and the ontology could consist only of the node ALL, with one gigantic script underneath it, containing all events as components and all objects as case-role fillers. Such superunification, while theoretically elegant, would not make the ontology particularly more useful however. Sound judgment must be exercised when deciding what becomes a script, a subscript, and a superscript. Sometimes it is very wise to make use of subscripts, in order to save time. Sometimes it is also very wise to expand an existing superscript, rather than creating an entirely new script. Quite often when this is appropriate, the events are conceptually close to each other. If anyone ever bothered to write a script for walking, for example, that person may as well expand it to running (and trotting, cantering, galloping, skipping, and jumping) because so much of the work involved in any one of these is relevant to the others also.

There has in the past been some discussion of implementing a more complete description of complex events, using logical operators and temporal markers; please see the references for a great deal of information about an experiment in doing this, including one very large experimental script for BANKRUPTCY. At the current time, however,

this has not been done and no particular plans have been made to implement these capacities in the system.

There is a slot that should not be confused with HAS-PARTS, or vice-versa, called HAS-OBJECT-AS-PART. This is, obviously, for describing the components of a complex object. Acquirers working without the benefits of KBAE's policing quite often try to use the HAS-PARTS slot for objects.

What to Acquire?

Knowing how to acquire items is unfortunately only half the battle; what to acquire is the other half and it is not easily fought or won. The problem of organizing acquirers and resources effectively existed while KBAE was usable, and is compounded now that it is not. In the very early days of acquisition, acquirers were told to add anything and everything they wanted, because the necessity to build a core world-representation was so desperate. Because of this, many of the "obvious" concepts and words have been acquired (the group of Russian acquirers found it difficult to think of a verb that was not already present in the English lexicon, for example). In spite of this good general coverage, it still becomes obvious that the system is lacking as soon as one tries to use it. Good acquisition strategies are needed.

Ontological and lexical acquisition are the meat and potatoes of Ontological Semantics. To write the ecological analyzer, parser, tokenizer, and TMR generator/producer are finite and relatively small tasks compared the acquisition of

mental concepts and of the words of a language. Much of the available resources have been devoted to acquisition of English words and the concepts necessary to represent them, and in the particular case of English, a person with high ambitions may declare that he will acquire the whole of the Oxford English Dictionary. Gigantic as this task is, however, it is not sufficient, because the OED is not a sufficient representation of English usage. New slang terms and technobabble are constantly being introduced, and the OED's account of compound words is often surprisingly insufficient, especially in light of the capabilities of computer processing as opposed to human processing. Upon hearing the term "fish fork," for example, a human assumes that it is a fork used to eat fish or possibly a fork made in the shape of a fish, and the compound is therefore absent from the OED. The text processing system, however, would need to be told these two possible meanings explicitly, as it is unable to rule out other impossible options such as a fork made out of chunk light albacore, which is only slightly more ridiculous an idea than a necklace made of candy or an airplane made of paper. There is also the huge problem of metaphor, examples of which may be found in the translations from *Cyrano de Bergerac* above. How does one think of all the possible metaphors for the meaning of a word? How does one see a scabbard as a scabbard, a sleeve, a phallic symbol—and who knows what else? The multiple senses of a word are commonly metaphors of each other, to a greater or lesser degree. The human mind excels at metaphorization in speech and most good dictionaries are therefore brimming with different senses of a word that differ from each other only in minute details. For instance, the entry for fork (noun) in the Oxford English Dictionary is as follows (excluding literary references, for the sake of relevance and brevity):

I. A pronged instrument.
1. a. An implement, chiefly agricultural, consisting of a long straight handle, furnished at the end with two or more prongs or tines, and used for carrying, digging, lifting, or throwing; also with word prefixed indicating its use, as digging-, dung-, hay-, etc. fork: see those words; also FIRE-FORK, PITCH-FORK, etc.
b. A similar implement used as a weapon.
c. The forked tongue (popularly supposed to be the sting) of a snake. Obs.

2. a. An instrument with two, three, or four prongs, used for holding the food while it is being cut, for conveying it to the mouth, and for other purposes at table or in cooking. For carving-, dessert-, fish-, pickle-, table-fork, etc. see those words.
b. forks and knives: the name of the club-moss Lycopodium clavatum.

3. a. Used in pl. for the prongs of a fork. Also transf. Cf. 12.
b. pl. (slang). The fingers. Hence, a pick-pocket (B. E. Dict. Cant. Crew ?a1700).

4. A steel instrument with two prongs which, when set in vibration, gives a musical note; called more fully a tuning-fork.

II. Applied to various objects having two (or more) branches.

5. a. A gallows. Also pl. Cf. FORCHE 1.
[So OF. fourche(s, L. furca; the Roman gallows was originally of the shape .]
b. Rom. Ant. Used to render L. furca, (a) the 'yoke' under which defeated enemies were made to pass as a token of their submission; (b) the forked stake used as a whipping-post.

6. A stake, staff, or stick with a forked end; a. as a prop for a vine or tree; b. a rest for a musket; cf. FORCAT. c. (See quot.). d. Mining (Derbysh.): see quot. 1881. e. A divining-rod.
a. 1389 Helmingham MS. 21. 17b, Forkis..to bere up e vyne. 1626 BACON Sylva §423 Some have put two little Forks about the bottom of their Trees, to keep them up~right. 1816 KEATINGE Trav. (1817) I. 43 The boughs..propped up by forks.
b. 1591 GARRARD Art Warre 7 To traine hys Forke or Staffe after hym whilest he..doth charge hys Musket.
c. 1726 Gentlem. Angler 149 A Fork. Vide Rest [for a fishing rod].
d. 1747 HOOSON Miner's Dict. Giijb, If..we think it will let the Forks settle when they come to be weighted, we put a Sill under them. 1881 RAYMOND Mining Gloss., Fork..a piece of wood supporting the side of an excavation in soft ground.
e. 1886 A. WINCHELL Walks & Talks Geol. Field 137 Some..even resorted to the witch-hazel fork [in 'prospecting' for petroleum].

7. Building. See quots. 1868, 1883.

8. Anat. fork of the throat or breast: app. the sternal bone together with the clavicles. Obs.

[= med.L. furcula, OF. fourcelle; the words seem to have been used very vaguely, and it is often impossible to determine the exact sense.]

9. The barbed head of an arrow. Obs.

10. In various technical uses. a. A piece of steel fitting into the socket or chuck of a lathe, used for carrying round the piece to be turned.
b. (also forks): see quot.
c. The front or back projection of a saddle.
d. The part of a bicycle frame in which the (front or back) wheel revolves; also attrib., as fork-blade, -crown, -end, -head. Also pl.

11. Mining. (See quots.)

12. [From the verb.] A forking, bifurcation, or division into branches; the point at which anything forks. Hence, each of the branches into which anything forks. a. gen.
b. In the human body, the part at which the lower limbs proceed from the trunk. Also (sing. and pl.), the lower limbs themselves; the lower half of the body. Cf. FORCHURE.
c. The point at which a river divides into two, or the point of junction of two rivers; a branch or tributary. Chiefly U.S.
d. of a road.
e. of a plant or tree.
f. A flash (of forked lightning); a tongue of flame.
g. Chess. A simultaneous attack on two pieces, esp. by a knight.

13. fig. a. nonce-use. The union of two lines of descent. b. A dilemma, choice of alternatives; also, a dichotomy, distinction. Obs.

14. Caudine Forks = L. Furcæ or Furculæ Caudinæ: proper name of a defile near Caudium, in Samnium, where the Romans were intercepted in the second Samnite war. Hence proverbially used for: A crushing defeat.

III. attrib. and Comb. (see also sense 10d).

15. a. objective, as fork-grinder, etc.; b. parasynthetic and similative, as fork-like, -shaped, -tongued adjs.; fork-wise adv.

16. a. Special comb.: fork-beam Naut. (see quot.); fork-beard, a name given to various fishes of the genus Phycis; fork-breakfast (see quot.); fork-carving a., that uses a fork in carving; fork-chuck (Wood-turning), a chuck with two or more teeth: see quot. 1874; fork-fish, ? a kind of thornback; fork-lift truck, etc. a vehicle fitted with a pronged device in front for lifting and carrying heavy goods; also fork-lift ellipt.; fork-moss, a kind of moss (Dicranum bryoïdes); fork-ribbed a., having ribs branching off like the prongs of a fork; fork-shaft, the handle of a fork; fork-staff-plane, a kind of joiner's plane used for

working convex cylindrical surfaces; fork-way, a point where two roads meet or diverge, a fork; fork-wrench (see quot.). Also FORK-HEAD, FORK-TAIL.
b. fork supper (also -buffet, -dinner, -lunch(eon, etc.), a meal served at a buffet, etc., consisting of food suitable for eating with fork alone, making the provision of set places at table unnecessary.

fork, n.
Add: [III.] [16.] [a.] forkball Baseball, a pitch in which the ball is held tightly with the thumb, index, and middle fingers spread wide apart, in order to make it fall down sharply or behave in an otherwise unpredictable manner; cf. split-fingered fastball s.v. *SPLIT ppl. a. 5 a.

fork, n.
Molecular Biol. More fully replication fork. A Y-shaped region where the strands of a duplex DNA molecule are separated during replication.

As the reader can see, all of these many definitions are, without exception, obvious metaphors of each other. Indeed if one took “objects” in the sense that it is used as a handle in the ontology, meaning anything that is neither an event nor a property, one could say that the only definition needed was that following Roman number II “Applied to various objects having two (or more) branches.” The point is that acquirers working straight through a dictionary are going to have to make sound judgments about what really constitutes a new lexical sense, in terms of the time and effort available. They must also be constantly referencing a thesaurus, because while a single lexical entry will have several different senses, there will be several different lexical entries with a sense that is basically the same as any one of those senses. It would be a waste of effort to not rapidly propagate a perfectly good sense, once it has been written.

This leads into the last issue that complicates acquisition via a dictionary, viz. the dozens of sets of dozens of entries which are ripe for rapid propagation but which are almost never close to each other alphabetically—such as flowers, medicines, musical

instruments, and any number of other categories. If acquirers are chained to an alphabetical list of items to acquire they will waste an enormous amount of effort in writing basically the same entry over and over again. In short, it is simply not practical restrict an acquirer to one section of a dictionary. Not only does it not result in full and useful coverage of the language (what percentage of the OED is obsolete or obscure?) but it creates an enormous waste of resources when lexical entries are not propagated because the acquirers are afraid of stepping on each others' toes. All acquirers must be in frequent communication with each other, and the databases they are referencing must be updated frequently. Ah, for the good old days of KBAE…

A slightly less ambitious person will not want to devote himself to acquisition of a language wholesale, but will instead be working towards a more specific goal within a language. These goals tend to be either the acquisition of a certain category of words, as in Televnaja's (2004) acquisition of phrasal verbs, or more commonly the acquisition of enough words and concepts to allow the system to competently process texts about a certain domain. Past domains of interest have included sports, information security (Malaya 2005, Krachina 2005), medicine (Triezenberg 2004, 2005), law, and finance. As a result all of these areas are well developed in both the ontology and the lexicon.

An acquirer interested in adding a grammatical category of words, such as Televnaja's phrasal verbs, faces two steps in her task, first to find a good method of computational representation within the system, and second to find a list of all the examples of this grammatical category, and acquire them.

The acquirer interested in domain acquisition finds himself making a rather more difficult journey. An acquirer who is undertaking a domain must first devote some time

to researching his subject, so that he can make informed decisions about what ontological concepts are necessary to support the lexicon. Indeed, it is best if, for each domain, an acquirer can be found who has a general interest in the area and therefore already knows something about it, and is interested enough to learn more and with a passion for detail. The longer a person has had to think about the domain and how the various objects and events in it interact with each other, the more successful will be the process of acquisition. After the domain has been thoroughly researched, the acquirer may attack it from two directions: top-down and bottom-up.

In a bottom-up domain acquisition, the acquirer begins by amassing a considerable corpus of texts in the domain. He then begins acquiring lexical items as he comes across them in the text. A piece of software has been written (by Evguenia Malaya) that will search a text and pick out words that are not yet in the lexicon, however the acquirer must be awake of all of the phrasals, noncompositional compound words, alternate senses, etc. which cause so many problems in so many ways. The process of acquiring a lexical item involves finding one ontological concept to attach it to and a number of slots to more narrowly describe it; as the acquirer works, it is more than likely that he will fail to find appropriate concepts, and need to acquire them. He does so. In bottom-up domain acquisition, ontological concepts are acquired in no particular order and strictly on a need-to-know basis.

A top-down domain acquisition begins by creating the ontology of the domain. The acquirer will most likely create whole branches of the ontology specifically to suit the domain, such as a branch of specialized social roles, a branch of artifacts, a branch of locations, and a branch of events. Top-down acquisition results in a much tidier and more

browseable ontology than does bottom-up, because as the acquirer works in the ontology he has the entire domain in his head—he has made a "road map" of it before he begins acquiring, with lists of what he wants to exist. He is much more likely to make the best possible use of inheritance, because he can create a master concept for each branch which contains all of the information that will be recycled in all of the child concepts. For instance, an acquirer working in the medical domain can make a concept for MEDICAL-PLACE, and in that concept note that it contains all of the pertinent social roles and artifacts, and that it is the LOCATION of all the appropriate events. The concepts for HOSPITAL, CLINIC, OFFICE, etc. will then all share this information.

Bottom-up acquisition produces solid results much more quickly than does top-down. Because the acquirer is focusing on actual texts in the domain, the words he acquires are the words that are actually used; and because he does not acquire a concept until it is needed, there is absolutely no waste of ontological acquisition. If time and money are tight, bottom-up acquisition is the way to go. It is more likely, however, to result in a messy ontology and to squander some effort in not effectively organizing branches for inheritance. The acquirer is more likely to decide he will "make do" with a concept that isn't quite right, than to create a new one that is—which can be either good or bad, depending on how urgently his project needs to be finished.

Top-down acquisition is structurally beautiful. The ontology branches in an orderly and coherent fashion, which makes it more pleasant to browse and increases the likelihood that other acquirers in the future will be able to find the concept they are looking for by sliding down the ontology. It is likely that concepts will be created that no lexical item will be linked to, at least for the time being—whether this is wasted effort or

not again depends on how urgent the work is; if a concept is not used now, it will be used in the future. The one major danger of top-down acquisition is that the acquirer will think too hard about what to acquire, try to be too clever, and end up spending a lot of effort creating whole branches that are completely unnecessary. This happened when the author acquired the medical domain, and made a whole branch for ANIMAL-PART, and then a whole branch for ANIMAL-DISEASE that was a virtual mirror of ANIMAL-PART. It was only after a couple hundred hours of lexical acquisition that the author realized she may as well have acquired one concept for ANIMAL-DISEASE, or possibly a handful of concepts for distinctly different disease processes, and used a slot to indicate what ANIMAL-PART the disease affected. This was a big loss indeed, but a valuable one, as the author may now alert others to such dangers. If an acquirer keeps his head and has prior experience that has led to a good understanding of the value of parsimony (which is computationally nil, but it was Nirenburg and Raskin's wish that the ontology remain as small as possible, and we respect that) and of the uses of slots, these blunders can be avoided. If a domain is being acquired purely as an academic or aesthetic (or Machiavellian) exercise, then top-down is the way to go, for the reasons listed above. If the domain is being acquired on a grant or for a project, then it is probably better to go with bottom-up and create a usable library of entries sooner.

So when do we start acquiring? Now. The author proposes to do a bottom-up acquisition of the domain of emotion in humans—to create an ontology that will support the lexicon of emotion words in the English language. Chapter 3 will focus on researching the domain, and chapter 4 will consist of the domain as acquired into the ontology.

Some domains are inherently organized, and the manner in which they should be ontologized is more or less obvious. An acquirer who is working on the domain of animates, for example, may simply construct the accepted taxonomy of Kingdom Animalia. He may acquire progressive levels of the ontological tree to mirror the Kingdom-Phylum-Class-Order-Family-Genus-Species progression of the taxonomic tree. Though in reality the taxonomy branches out in many more levels, this restriction is an appropriate exercise in parsimony; so too would be excluding those branches of all levels that accommodate animals both extinct and generally unknown. In some branches, it would be appropriate to reserve the level of Species or even of Genus for the lexicon (though for the example of domestic dogs, *Canis familiaris*, one wants both in the ontology—breeds go in the lexicon).

The best structures of other areas are less clear, for example the medical domain. It needs to cover anatomy and physiology, pathology and normal process, kinds of treatments, special tools and places, as well as all the other associated domains mentioned above. The question becomes even more interesting when the domain in question is subjective—when it is a mental or even social domain, rather than a concrete. How does one discover the structure of such a domain; how does one gather together all of the necessary ideas, actions, and objects associated with it; and the most daunting questions of all—should one manage to do so, to what degree is it an exercise in emics rather than etics, how would other people have structured the domain, can only one structure of the domain adequately process the semantics of any given text (which is the main point, after all), if it can't then how many different structures need to be added, and once all of these

structures are added, will any structure remain at all, or will the domain just be one gigantic conceptual web with no form or meaning whatsoever?

This project is to outline heuristics and methodologies for domain acquisition, and to wade into the sticky waters of subjectivity—into the domain of *emotions*. Onward to chapter 3, then.

CHAPTER 3: HEURISTICS

The first step in this project is to see what already exists in the ontology by way of emotion concepts. We find a whole branch called EMOTIONAL-EVENT, which is exactly what we were looking for. The branch itself, however, is a mess. It is roughly divided into POSITIVE-STATE and NEGATIVE-STATE, but with a few other concepts that would seem to belong under one of those two also located in the main branch. In several places, a parent concept has only one child concept, which is definitely not acceptable. In addition, the idea of "emotion" seems to have spilled over into events that may excite or result from emotion, examples being LAUGH, MOCK, INTIMIDATE, and SHED-TEAR. As inelegant as the title of the last concept may be (why not CRY or WEEP?) it cannot be changed: remember that version control without KBAE requires that no concept ever be deleted or re-named, so that other acquirers' work as well as lexical linking is not spoiled. It is most definitely the intention of this project to raze the branch down to EMOTIONAL-EVENT and then rebuild it entirely; whatever ontology is built to replace what exists, though, must somehow accommodate all of the legacy concepts. We will keep this in mind as we design the new branch. The complete legacy branch may be perused in Appendix B.

The first step, in delimiting a subjective domain, would seem to be finding what part of it is generally agreed upon—does it have any basis in scientific knowledge; is

there a field or subfield dedicated to its study, that has gone before you and found or imposed a standard structure upon it; at the very least, what are the lexicon items one will want to have present in the system? These three approaches yield the following results, for emotions.

1) A basis in scientific knowledge: it is known that certain changes in the brain are associated with certain emotional states. Chemical depression, for example, is associated with an excess of dopamine. A shock to the amygdala causes a cat to hiss and spit, so activity in that region is associated with whatever induces an animal to fight (Kaada 1967; Panksepp 1982). The autonomic nervous system, too, can bring about two distinct states in an organism: sympathetic activity causes arousal, and parasympathetic activity induces calm. Both of these have certain neurochemicals and hormones as well as certain observable effects associated with them (Ellsworth 1994; but also Cacioppo *et al*. 1993; Kagan 1997). Any number of words may be connected with the state of sympathetic arousal: panic, fear, anger, anxiety, and desperation, to name a few. The fact that so many words, which have little in common semantically, are associated with this state reveals that its grain size is not small enough for the project at hand. It may well prove useful to have a branch of emotions associated with sympathetic activity and another with parasympathetic, but more branches, both at the same level and at lower, are needed.
2) The study of emotion has been of particular interest to psychologists (and to students in psychology classes). This study, as opposed to those in (a) which are more likely to be conducted by neurobiologists, has occupied itself with creating

theories of how emotions are "structured" with relation to each other and to behavior—their manifestation. Research that has focused on the relation between perceived emotion and manifested emotion—that is, what people report feeling and their correlated behavior—has been unable to resolve the issue of whether the behavior associated with emotions (autonomic, pheromonal, socially conditioned, or otherwise) creates the experience of emotion, or whether experience of emotion causes the behavior (Lang 1994; Levenson 1992; Cannon 1927). A chain of causality in the ontology of emotion would, therefore, be inappropriate—though emotions and behaviors should definitely be linked. Other studies have shown that certain facial expressions are recognized around the world (Ekman 1981; Ekman *et al*. 1983, 1983; Keltner and Buswell 1996; but also see Russell 1994). This is a promise, to the lone acquirer, that some common ground may be found in all personal ontologies. Lastly, some research has theorized about possible hierarchies or categories of emotion (Oatley and Jenkins 1996; Plutchik 1994): these theories have the ontologist's interest very much at heart. These sets of categories seek, as all theories and surmises do, to explain as much as possible with as little elaboration or complication as possible. Parsimony is therefore built into them, and the acquirer should test the corpus against them, first to discover if there are categories which are unused (in which case, the acquirer should stop and think, hard, about what he or she may be missing), and second to discover if the presented categories are sufficient (in which case again, the acquirer should stop and think hard, this time about how many concepts are really needed, and how

they can be nuanced through the use of properties, both in the ontology and in the lexicon). Acquiring a corpus leads into the third avenue of investigation,

3) Examining a corpus in order to discover what words are used, and in what ways they are used. To settle on an appropriate corpus for such an area as emotions will be extremely difficult. The most appropriate corpus, in an ideal world, would be all of literature, from all places and all times—a wildly inappropriate approach, in the real world. While it is true that the meaning of everything eventually spills into the meaning of everything else, not every domain's proper corpus is quite so diffuse: the medical corpus could call itself complete if it contained a manual of diagnosis and treatment, a handbook of pharmacology, a medical dictionary, and some patient data (the collection of which would be a research problem in itself, the most feasible source being public message boards about health). Probably the best place to start, when creating a list of words and their senses for the domain of emotion, would be a very good dictionary such as the OED. The acquirer could pick any term related to emotion, lexicalize it according to the broad and contemporary usages, and from the word's entry (which would be extensive) pick out other words relating to emotion. In this manner, the acquirer might work through fine gradations of meaning, collecting an exhaustive list of words and at the same time testing the structure of the ontology. Do the tints and shades of meaning found in the literary samples of OED entries follow the tints and shades allowed for in the ontology? If not, it needs to be rethought. In this avenue especially, the acquirer must be flexible: the ontology is constantly growing and being revised, and so must the acquirer extend and revise his own area of it. It is

> vital that the acquirer not be paralyzed by indecision, or the feeling that a certain amount of information needs to be accumulated before acquisition, either ontological or lexical, can begin. In this, as in many things, the best course of action is to simply plunge into the work, knowing beforehand that any entries which may be discarded are not lost effort, but a sign that the accuracy and scope of his work will be all the better in the end.

As promised in the previous chapter, the acquisition of this domain will be top-down rather than bottom-up. This is partially because the acquisition is an academic exercise rather than a practical one—no one is interested in developing the "emotions" branch of the ontology for applications starting next Tuesday. The development may therefore proceed at a leisurely pace and with the aim of creating a high-quality product. The substantive reason for top-down acquisition, though, is that point (3) is not feasible for this domain—and that is the whole point of this project. It would be easy and, at some grain sizes justified, to acquire a new concept for every emotion-related word in the corpus. They all have unique shades of meaning, after all. Doing so would not, however, result in any kind of order in the ontology, and would in fact result in chaos: hundreds and hundreds of unorganized sibling concepts, unbrowsable, unusable, and certainly not well defined. This domain is too squishy to be acquired bottom-up. It needs a backbone. In other words: reference to the lexicon will be nearly or completely irrelevant in this domain's acquisition, because we plan on creating an ontology that will be defined as complete before we ever begin to examine texts. The lexical acquisition phase would be an exercise in pounding square pegs into round holes; all emotion words simply must be made to fit into what exists. It will not be a business of adding and refining the ontology

to suit whatever comes up—unless what comes up is very seriously underrepresented indeed. We therefore rely heavily on the psychologists to have done their work well and thoroughly, so that our ontology can bear its burden.

At this point, it would be prudent to repeat—to stress—that this project is in no way interested in finding out "the truth" about human emotion. This is not a dissertation in psychology and the author is not a psychologist, an undergraduate minor in psychology notwithstanding; no experiments have been run to determine what emotions "really do" exist, or whether any exist at all, or if they do exist, if there is such a thing as a set of primary or basic emotions. This project is to create a smallish core of concepts that can be manipulated to adequately describe whatever lexical item it may be required to interpret. This does create a labor loop: we want an ontology that adequately describes the emotions lexicon, but we are going to make the emotions lexicon fit the ontology, but we want the ontology to adequately describe the lexicon… such is the problem of all ontological efforts outside of the very strictly delimited and hierarchical domains such as plants and animals. We want to create databases that understand the world—but the only world the databases will ever understand is the world we describe to it. This is why the issue of multiculturalism in the ontology is such a hot-button topic, why it is necessary to have many acquirers from many backgrounds work in the ontology, and why a multiple inheritance structure rather than a strict taxonomy is a must. At the end of the day, though, the consumer must accept that the ontology will never be more than **good enough**. With no explicit, divine truth to be uncovered, all that is left are fragments of human experience and a system that is **good enough** to handle it.

Psychology textbooks frequently pair the discussion of emotion with the discussion of motivation, perhaps because they are intertwined (one feels emotion according to the satisfaction or frustration of one's motivation) or perhaps because it is comfortable to say that motivation is shared by all animals including humans, while only humans can be definitively said to have emotions. There are at the very least four motivations: fight, flight, feeding, and sex (which can be made to alliterate, but not in dissertations). It can be argued that there are many more, especially for our own species—Maslow's (1954) hierarchy of needs, for example, could be called a hierarchy of motivation. Though psychology textbooks consistently lump emotion and motivation into the same chapter or section of the book, though, they do draw a line between the two. It is always emotion *and* motivation. We will therefore focus our energies on emotion only, perhaps leaving motivation for a later project.

Because the ontology is supposed to be "definitive," it is probably best to play it safe by combining as many different theories of emotion as possible, both those which posit a set of primary emotions and those that don't. In a very handy review of primary-emotion theories, Ortony and Turner (1990) gathers and critiques the major players in the field at the time. Meekly assuming that the range of human emotion has not changed significantly in the fifteen years since this review was published, let us begin with their data. Here is a compendium of all the proposed basic emotions from Table 1 of this paper:

> Acceptance, anger, anticipation, anxiety, aversion, contempt, courage, dejection, desire, despair, disgust, distress, elation, expectancy, fear, grief, guilt, happiness,

hate, hope, interest, joy, love, pain, panic, pleasure, rage, sadness, shame, sorrow, subjection, surprise, tender-emotion, terror, wonder

Plutchik's (1994) theory of emotion is actually more comprehensive than the table in Ortony and Turner's article gives it credit for. Each of the eight emotions comes in gradations of strength, which Plutchik has given names to. Also, he names eight intermediary or secondary emotions. He organizes all of these emotions into an "emotional solid," shaped approximately like a cone, with the "point of emotional zero" at the bottom. The emotions grow stronger as one moves upward and the cone expands, and one moves through the primary emotions, via the secondary emotions, as one rounds the circumference. The complete collection of named emotions in this solid are: (serenity, joy, ecstasy) love (acceptance, trust, adoration) submission (apprehension, fear, terror) awe (distraction, amazement, surprise) disapproval (pensiveness, sadness, grief) remorse (boredom, disgust, loathing) contempt (annoyance, anger, rage) aggressiveness (interest, anticipation, vigilance) optimism

With a little semantic tinkering, one can say that this list comprises the list above, with the exception of pleasure and pain—two items that this study of neuroscience is perfectly prepared to call sensations rather than emotions. Comparing the two lists certainly involves a little fudging; are wonder and awe really the same thing? Are love and tenderness? Are anxiety and apprehension? Is shame a kind of remorse, or rather disgust for oneself? And indeed, some of the items Plutchik has called "emotions" seem to be something else. What about disapproval? Is it really an emotion or is it a cognitive construct? Ortony and Turner argue that several of the so-called emotions are really cognitive states, with surprise as a special example. They assume that an emotion is either

positive or negative, an assumption that he himself acknowledges is weak, and that surprise may be a reaction to either a positive thing or a negative thing, and is therefore not an emotion but a cognitive state. I disagree with the idea that an emotion is positive or negative, on the grounds that it is an unfounded value judgment and that sadness and anger are sometimes positive experiences.

I would also like to argue that emotions in general are not separable from cognitive states. One expects (as a matter of general experience) an angry person to be cognitively different from a happy one or a sad one. An angry person will be aggressive and energetic; a sad person will be reluctant to act; a happy person will agree to things that, in other states, he normally wouldn't. Some emotions are states of physiological arousal and some aren't. An emotion is really a state of altered cognition and Ortony and Turner's argument against surprise does not stand. Perhaps on these grounds my own argument against pain and pleasure does not stand, for one expects people experiencing these sensations to react very differently. I think, though, that these two particular concepts are too tightly wound up with motivation to be appropriate to emotion—though it is quite possible that when motivation is added to the ontology, pain and pleasure will be a common bond between them.

Some emotion theorists see emotions as tripartite experiences (Nairne 1997): a physiological reaction such as arousal, an expressive reaction, and a subjective experience. Quite arguably, cognitive states (whatever they are) consist of the same three parts—a physiological reaction in terms of the neurotransmitter balance in the brain, an expressive reaction in terms of how the person will react differently to the same situation when in different states, and the subjective experience of the thought processes associated

with the state. In short, the distinction between emotions and cognitive states does not seem to be a useful one.

Let us begin, then, by taking Plutchik's model as the starting place for our ontology of emotions. There are eight primary emotions which can be the first level beneath EMOTIONAL-EVENT, and eight secondary emotions which can be the level below that, with each secondary emotion being a child of the two primary emotions it is (supposedly) a mix of. There is a property called INTENSITY which currently has in its domain PHYSICAL-EVENT and SOCIAL-EVENT. It would be reasonable to add to this domain, I think, MENTAL-EVENT (and probably equally reasonable to remove SOCIAL-EVENT; it is not clear what it is doing there). The property can then be used in lexical items to define the intensity of the emotion, which will accommodate Plutchik's entire emotions model. In addition, the lexical items should have either an AGENT or an EXPERIENCER that is set by default to HUMAN, and a THEME which can be set to ALL by default but made more specific when appropriate—for example, "shame" can be linked to DISGUST with a THEME that refers back to the EXPERIENCER.

The notion of expressive reactions as an important component of emotions is an interesting one, and it is precisely the kind of information that the ontological semantic system could find useful. Texts very often attempt to describe the emotional state of a person entirely through that person's actions. This is prevalent in fiction and the system is less likely to deal with fiction that with non-fiction, but there are a lot of non-fiction texts that would do this also, such as news reports and annotated interviews. Having decided that it will be useful to acquire a few actions that stereotypically correlate with certain emotions, how do we decide which actions to acquire and which emotions to associate

them with? Ekman's (1977, 1981, 1992; see also Ekman *et al*. 1983) work with facial expressions is a good starting point: he has tested the facial expression/emotion question in cultures around the world and found a good correlation in several cases. We will look at his research, find out what the elements of stereotypical facial expressions are, and include them. Ekman is happy, in his work, to publish photographs of faces representing each of his basic emotions, and also happy to give one a chart of which facial muscles are involved in expressions (Ekman 1977). He is a little less clear about which muscle movements correspond to which emotions. Ekman (1992) does correlate a few; the movements correlated in this text are as follows:

1) Eye movements
    a. Widening eyes
    b. Narrowing eyes
    c. Gaze up
    d. Gaze down
    e. Blink
    f. SHED-TEAR
2) The visible autonomic responses
    a. Blushing
    b. Blanching
    c. Sweating
3) The brow
    a. Triangulation
    b. Drawing or lowering

   c. Raising

4) The mouth

   a. Narrowing

   b. Zygomatic smile (positive emotion)

   c. Risorious smile (fear)

   d. Contempt smile

   e. Dampened/miserable smile

Also to be considered are the symptoms caused by action of the sympathetic and the parasympathetic nervous system—that is, typical signs of arousal and of relaxation. While some emotions are not clearly linked with either (happiness and sadness, for example, are really neither here nor there) some clearly are (such as fear). The standard symptoms of arousal should therefore be linked to some emotional concepts.

The acquirer may also turn to his own folk knowledge in this situation, because the stereotypical ideas of what an emotion looks like must certainly influence textual representation of emotions, whether the ideas have any relation to reality or not. The acquirer will therefore acquire a handful of what she considers to be stereotypical actions representing each emotion. Future acquirers working in this branch are of course expected to add their own ideas.

Here then is the proposed map of the ontological concept for emotion:

1) Primary emotions seated directly beneath EMOTIONAL-EVENT and with two child concepts each
2) Secondary emotions seated beneath two primary emotions
3) A HAS-PARTS slot which includes:

a. Physiological components of the emotion
b. "Universal" physical components of the emotion
c. Physical components of the emotion in the acquirer's script for the emotion

Is it possible that it would be appropriate to include concepts for SYMPATHETIC-REACTION and PARASYMPATHETIC-REACTION underneath the concept EMOTIONAL-EVENT? Here is the acquirer's instinct: sympathetic and parasympathetic responses stem from parts of the neural system below the brain, that is, one doesn't have to have any subjective experience of arousal or relaxation in order for the body to be actually aroused or relaxed. Indeed, it is standard for people who have undergone a major stressor to not realize that they are having a sympathetic response until after the stressor has been dealt with. For this reason, sympathetic and parasympathetic responses do not belong in the EMOTIONAL-EVENT branch. Also, the ontology relies heavily on "folk knowledge," and in the acquirer's world (which is the only world which she can be expected to acquire into the ontology) there are people who do not believe that creatures other than humans can experience emotion—or that creatures other than large animals can experience emotion—and so on up the phylogenetic tree. Whatever anyone believes, it is true that physiological arousal is a state experienced by almost all complex animals. This is not as strong an argument for not including sympathetic response in the emotions, but it supports the argument previously given. Another piece of folk knowledge is that a person who is truly undergoing a sympathetic response feels devoid of emotion while the response is in effect; a person who is actually in danger reacts coolly and quickly to what is going on. Emotion is then

experienced as the parasympathetic response is activated and the person is again able to wonder “what if” and “should I.”

Another plan of attack is not evidence that the physiological responses should not be emotions, but rather a lack of evidence that they should: emotions theorists in the late nineteenth and twentieth centuries spent a lot of time trying to decide if the subjective experience of an emotion precedes or follows its physiological and physical components (James 1890, 1894; Cannon 1927; Schachter and Singer 1962). Argument over the temporal order of emotions components has fizzled out over time, though; for whatever reason it has ceased to be of interest. The psychologists have spoken… or failed to speak.

If sympathetic and parasympathetic responses are put elsewhere in the ontology they can still be used as fillers in the emotions concepts, acknowledging that the two are intertwined, but also acknowledging that they are separate sorts of experiences. Creating these two concepts will make the emotions concepts much cleaner and save a lot of work in acquisition: not only can all of the components of the sympathetic and parasympathetic responses be grouped into two concepts, but these components no longer need to be concepts in their own right. They can be described using slots. These two concepts also cannot fail to be useful to many other acquirers working in other areas. Therefore, we will create them.

Contrary to the manner in which we have decided to handle physiological responses, there is evidence (Tomkins 1962; Matsumoto 1987; Eibl-Eibesfeldt 1970; Izard 1994) that facial responses are very closely tied to the subjective experience of emotion—Matsumoto and Tomkins both propose that simply making the face associated with an emotion can influence a person’s mood, and therefore recommend that people

suffering from clinical depression force themselves to smile often. Whether this is a valid therapy or not, it does seem that facial expression is quite closely tied to the subjective experience of emotion. It is also species-specific. A chimpanzee that smiles is not happy—and only humans have been observed to SHED-TEAR, in terms of the full theatric display of weeping. It therefore seems mete to add another concept directly beneath EMOTIONAL-EVENT entitled something like EMOTIONAL-RESPONSE, where both facial responses and stereotyped actions (things that are not facial expressions, such as raising the voice, weeping, and laughing) can be contained.

Laughter creates another problem for us. Is the experience of finding something funny an emotion? Is laughter an emotional response? It seems clear that it must be, though laughing and crying together are basically unique to the human species and troublesomely difficult to explain away in Darwinist terms. One can cry when one is angry, afraid, sad, grieving, and happy. One can laugh when one finds something funny, when one is triumphant, or from sheer joy. Plutchik's model includes joy, but it most definitely does not include humor and to accommodate triumph we will have to start out with joy and tweak its SOURCE (triumph is a good example of an emotion that is the direct result of satisfying a motivation). This particular acquirer has published in the area of humor and listened to many talks about all aspects of it—it therefore takes a primary place in this acquirer's concerns, and she feels that it deserves a concept of its own. It is most definitely not the same thing as joy. So, there will be a freestanding concept beneath EMOTIONAL-EVENT, separate from the oh-so-neatly made Plutchikian daisy-chain.

With this start under our belts, let us begin acquisition and see what other issues evolve.

CHAPTER 4: THE PROPOSED ONTOLOGY

The Blueprint

As this ontology is being acquired without the benefit of any tools that display its structure in a graphical, browseable format, it is useful—necessary—to maintain a "blueprint" of the domain. A blueprint is a list of all the concepts in the domain, with successive indentation indicating the parent/child relationships between concepts. This blueprint helps the acquirer to keep track of what has been acquired and what has not (as much to prevent concepts from being acquired more than once, as to make sure that none are missed), to browse the list of concepts as it grows into multiple pages, and to keep track of what needs to be in the IS-A and SUBCLASSES slots of each concept. The blueprint of the proposed domain of emotion will be as follows:

ANIMAL-LIVING-EVENT
    autonomic-response
        sympathetic-response
        parasympathetic-response

EMOTIONAL-EVENT
    HAPPINESS
        optimism
        love
            APPRECIATE
    TRUST
        love
            APPRECIATE
        submission
    HAVE-FEAR

submission
awe
WORRY
SURPRISE
awe
disapproval
FASCINATE
SADDEN
disapproval
REGRET
SHAME
PITY
GRIEVE
DISGUST
REGRET
SHAME
SCORN
HATE
ANGER
SCORN
ENVY
RESENT
HATE
FRUSTRATE
anticipation
ENVY
RESENT
optimism
FRUSTRATE
GRIEVE
WORRY
FASCINATE
humor
EXCITE
NEGATIVE-STATE
POSITIVE-STATE

emotional-action
frown
narrow-mouth
smile-zygomatic
smile-risorious
smile-contempt
smile-miserable
draw-brow

furrow-brow
raise-brow
blush
blanch
LAUGH
SHED-TEAR

ABSTRACT-SOCIAL-ACTIVITY
CONSOLE-EVENT
INSULT
DEMORALIZE
INTIMIDATE

EXTERNAL-VERTEBRATE-PART
eyebrow

OPPOSITION-EVENT
MOCK

POLITICAL-EVENT
AGITATE

PASSIVE-COGNITIVE-EVENT
SUFFER
ENJOY
COMPOSURE

UNDERSTAND
SYMPATHY

INVOLUNTARY-PERCEPTUAL-EVENT
EMPATHY

DESIRE
YEARN-FOR
FEEL-NOSTALGIC

REMEMBER
FEEL-NOSTALGIC
GRIEVE

The New Material

Here begins the real work of the domain: the acquisition of new concepts and the rethinking of legacy concepts. Throughout this section, the parts of legacy concepts that have been preserved are in CAPITAL LETTERS and any new information is in lowercase.

```
(autonomic-response
(definition (value (“a response of the autonomic nervous sytem”)))
(is-a (value (animal-living-event)))
(subclasses (value (sympathetic-response parasympathetic-response)))
)

(sympathetic-response
(definition (value (“the sympathetic response of the autonomic nervous system”)))
(is-a (value (autonomic-response)))
(has-event-as-part (sem (
        increase (theme (sem (pulse body-temperature blood-pressure)))
        sweat
        blush
        blanch
        decrease (theme (sem digest)))
        change-location (theme (sem (blood))) (source (sem (skin digestive-system)))
        (destination (sem (brain muscle)))
        loosen (theme (sem (lung)))
        produce (theme (sem (hormone)))
        excite
)))
)

(parasympathetic-response
(definition (value (“the parasympathetic response of the autonomic nervous system”)))
(is-a (value (autonomic-response)))
(has-event-as-part (sem (
        decrease (theme (sem (pulse body-temperature blood-pressure)))
        increase (theme (sem digest)))
        tighten (theme (sem (lung)))
        produce (theme (sem (hormone)))
        relaxation
        blush
)))
)
```

```
(HAPPINESS
(IS-A (VALUE (emotional-event)))
(DEFINITION (VALUE ("having, showing, or causing great pleasure or joy")))
(subclasses (value (optimism love)))
(has-event-as-part (sem (
        parasympathetic-response
        zygomatic-smile
        shed-tear)))
)
```

This is one of Plutchik's eight primary emotions. In Plutchik's model it represents the "joy" dimension, but I have re-used the concept HAPPINESS to accommodate the legacy ontology. I have included one autonomic response sub-event, one facial sub-event, and one stereotyped action sub-event (smiling would of course have been the first choice for this, but it had been taken care of, so shed-tear was elected to represent the fact that people sometimes cry with joy).

(TRUST
(definition (value (“the feeling of being able to rely on something”)))
(is-a (value (emotional-event)))
(subclasses (value (love submission)))
(has-event-as-part (sem (parasympathetic-response)) (not (worry)))
)

This is one of Plutchik’s eight primary emotions. I did not identify any concepts in the legacy ontology that it would replace. It was tempting to use the dot notation to indicate that one can only trust something outside of oneself, but on reflect, this is incorrect. One can trust oneself also, I trust.

```
(HAVE-FEAR
(SUBCLASSES (VALUE (INTIMIDATE)))
(IS-A(VALUE (emotional-event)))
(DEFINITION (VALUE ("to feel afraid")))
(subclasses (value (submission awe worry)))
(has-event-as-part (sem (
        sympathetic-response
        scream
        risorious-smile
        shed-tear)))

)
```

This is one of Plutchik's eight primary emotions. Instead of calling it just "fear," I have re-used the concept HAVE-FEAR to accommodate the legacy ontology. I have removed a useless slot (AGENT (SEM (*NOTHING*))).

```
(SURPRISE
(IS-A (VALUE (EMOTIONAL-EVENT)))
(DEFINITION (VALUE ("to come upon suddenly or unexpectedly- to take unawares")))
(AGENT (SEM (ANIMAL)))
(subclasses (value (awe disapproval fascinate)))
(has-event-as-part (sem (
        sympathetic-response
        scream
        raise-brow)))
(precondition (not (predict)))
)
```

This is one of Plutchik's eight primary emotions and also a concept from the legacy ontology. The handles were the same, so nothing had to be renamed.

```
(SADDEN
(IS-A (VALUE (emotional-event)))
(DEFINITION (VALUE ("to make or become sad")))
(subclasses (value (disapproval regret pity grieve)))
(has-event-as-part (sem (
        shed-tear
        draw-brow
        frown)))

)
```

This is one of Plutchik's eight primary emotions. Plutchik called it "sadness," but I have used the concept SADDEN from the legacy ontology.

```
(DISGUST
(definition (value (“to find something repulsive or intolerable”)))
(is-a (value (emotional-event)))
(subclasses (value (regret scorn hate)))
(has-event-as-part (sem (
        smile-contempt
        opposition-event)))
)
```

This is one of Plutchik’s eight primary emotions. Though there were extant concepts that were something like disgust, they were all actually disgust blended with something else. So, I have created an entirely new concept.

```
(ANGER
(IS-A (VALUE (emotional-event)))
(DEFINITION (VALUE ("to create or have hostile feelings because of opposition, hurt,
etc.")))
(subclasses (value (scorn envy hate frustrate)))
(has-event-as-part (sem (
        sympathetic-response
        smile-risorious
         furrow-brow
        frown
        opposition-event
        shed-tear)))
)
```

This is one of Plutchik's eight primary emotions and also a concept from the legacy ontology.

```
(anticipation
(definition (value (“to be in a state of expectation”)))
(is-a (value (emotional-event)))
(subclasses (value (envy optimism frustrate grieve worry fascinate)))
(has-event-as-part (sem (
        excite
        raise-brow
        sympathetic-response)))
)
```

This is one of Plutchik’s eight primary emotions. In Plutchik’s model, this dimension includes “interest” at the low-intensity end. This made me wonder if the legacy concept FASCINATE should replace ANTICIPATION; on reflection however I decided that they are not the same.

(optimism
(definition (value (“happy expectation”)))
(is-a (value (happiness anticipation)))
)

This is one of Plutchik’s eight secondary emotions. There was no concept in the legacy ontology that could be used for it, so it was created new. Note that this concept contains no new slots of its own. This is acceptable because it has inherited a unique combination of slots from its twin ancestry, which sufficiently differentiate it from other concepts.

(LOVE
(THEME (SEM (OBJECT)))
(IS-A (VALUE (happiness trust)))
(subclasses (value (appreciate)))
(DEFINITION (VALUE ("to feel strong affection for")))
(has-event-as-part (sem (protect help interact-socially)))
(intensity (sem (>.5)))
)

This is one of Plutchik's eight secondary emotions, and already existed in the ontology. I have included the slot intensity in order to differentiate it from two other extant concepts, APPRECIATE and ENJOY. This branch is begging to be further elaborated, perhaps on the grounds of C. S. Lewis' *The Four Loves* or just on the basis of what various objects and persons can be the object of it—love of self, of parents, of children, of brothers, of friends, of country, of home, of raspberry jam. This would be more an exercise in philosophy and sociology than in the study of emotion, however, so I have left it for later.

(submission
(definition (value (“to bow to the will of something”)))
(is-a (value (trust have-fear)))
(has-event-as-part (not (opposition-event)) (sem (control-event)) (relaxable-to (cooperative-event)))
(beneficiary (not (submission.agent)))
)

This is one of Plutchik’s eight secondary emotions. Nothing in the legacy ontology was like it, so it was created new. The main point of interest in this concept is the dot notation in the beneficiary slot; this is extremely useful when one wants to rule out everything except whatever fills a particular case role. I have indicated that one can only submit to something other than oneself. I justify this in terms of submitting to one’s own motivations and emotions on the grounds of work presented in my M. A. thesis, chapter 2, which establishes that a certain category of emotion has historically been interpreted as experience thrust upon a person, often again the person’s own will. Though the source of these passions is technically within the person himself, he has no agency over them and therefore the beneficiary of submission is not submission.agent.

(awe
(definition (value (“fearful wonder”)))
(is-a (value (have-fear surprise)))
)

This is one of Plutchik’s eight secondary emotions. Nothing in the existing ontology was like it, so it was created new. Once again it has no new slots but instead relies on having a unique combination of descriptors, derived from its multiple ancestry. I did not see any reason to restrict the beneficiary of this emotion to not awe.agent, on the grounds that one is sometimes awed by one’s own agency.

(disapproval
(definition (value (“to be surprised and saddened”)))
(is-a (value (sadden surprise)))
)

This is one of Plutchik’s secondary emotions. As with so many of the secondary emotions, there was no concept in the legacy ontology that adequately described it, so it was created new.

```
(REGRET
(IS-A (VALUE (sadden disgust)))
(subclasses (value (shame)))
(DEFINITION (VALUE ("to feel sorry about an event, one's acts, etc.")))
(precondition (sem (event)))
(beneficiary (sem (regret.agent)))
)
```

In Plutchik's model this is a secondary emotion called "remorse." The existing concept REGRET seemed to be the same thing, so it has been re-used. I have expressed the temporal nature of regret using the slot precondition; my reasoning is that regret always involves an event in the past—though sometimes the past event is a decision or realization about the future. I have also indicated, using dot notation, that one is in some way responsible for the things one regrets. This may have been bad judgment on my part, and might be changed in the future.

```
(SCORN
 (IS-A (VALUE (disgust anger)))
(DEFINITION (VALUE ("to regard or refuse with extreme contempt")))
(intensity (value (=<.5)))
(beneficiary (not (scorn.agent)))
)
```

This is Plutchik's secondary emotion "contempt," for which the existing concept SCORN has been used. In order to make peace with another legacy concept HATE, which seems to have much in common with scorn, I have differentiated them using the intensity slot.

```
(ENVY
(IS-A (VALUE (anger anticipation)))
(subclasses (value (resent)))
(DEFINITION (VALUE ("to feel discontent over or desire for other people's possessions,
accomplishments, etc.")))
(beneficiary (sem (human)))
(theme (sem (
        object (owned-by (not (envy.agent)))
        event (experiencer (not (envy.agent)))
)
```

According to Plutchik, aggressiveness is a mixture of anger and "anticipation," which has been equated by other researchers with "desire." Anger and desire are defining elements of envy, so I have used envy, an extant concept, in its place. Given that the theme of envy can be any object or event, I have chosen not to specify.

(emotional-action
(definition (value ("a physical action related to an emotion")))
(is-a (value (emotional-event make-gesture)))
(subclasses (value (frown smile-zygomatic narrow-mouth smile-risorious smile-contempt smile-miserable draw-brow furrow-brow raise-brow blush blanch laugh shed-tear)))
(part-of-event (sem (emotional-event)))
)

I have created this branch beneath emotional-event to include the observable reactions associated with emotions. This branch is a prime candidate to be developed into a very small grainsize. If the ontology had concepts for specific muscles rather than just for "muscle," for instance, then many of the emotional actions could be described minutely. Areas of the brain involved in these actions, neurotransmitters, and hormones could also be included (if the infrastructure allowed—and such detailed infrastructure was not required for the gigantic and highly elaborated medical domain, interestingly enough).

```
(eyebrow
(definition (value (“the brow ridges of the face”)))
(is-a (value (external-vertebrate-part)))
(location (sem (forehead)))
(above (sem (eyelid)))
(has-object-as-part (sem (skin muscle hair)))
)

(raise
(definition (value (“to physically move something to a place above its previous place”)))
(is-a (value (change-location)))
)

(lower
(definition (value (“to physically move something to a place below its previous place”)))
(us-a (value (change-location)))
)
```

An excellent example of how no domain is an island: neither the legacy ontology nor the anatomy branch of my own medical ontology had the concepts necessary to describe brow motions, so they had to be added. I placed “eyebrow” under external-vertebrate-part, where the concepts face, ear, and forehead also live. “Raise” and “lower” were created as empty children of change-location, a branch of the ontology which warrants an eventual overhaul and re-thinking. I have regretfully left their elaboration for later.

```
(draw-brow
(definition (value (“to draw the inner corners of the brows up and together”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem(
        tighten (theme (sem (eyebrow)))
        raise (theme (sem (eyebrow)))
        wrinkle (theme (sem (forehead)))
)))
)

(furrow-brow
(definition (value (“to draw the inner corners of the eyebrows together and down”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        tighten (theme (sem (eyebrow)))
        lower (theme (sem (eyebrow)))
        wrinkle (theme (sem (forehead)))
)))
)

(raise-brow
(definition (value (“to raise the eyebrows”)))
(is-a (value (emotional-event)))
(has-event-as-part (sem (
        raise (theme (sem (eyebrow)))
        wrinkle (theme (sem (forehead)))
)))
)

(frown
(definition (value (“to pull the corners of the mouth down”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (twist (theme (sem (mouth))))))
)
```

```
(narrow-mouth
(definition (value (“to press the lips together”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        tighten (theme (sem lip)))
        shrink (theme (sem lip)))
)))
)

(smile-zygomatic
(definition (value (“to smile using the zygomatic muscles, which creases the eyes and
raises the corners of the mouth”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        raise (theme (sem (mouth)))
        wrinkle (theme (sem cheek eye)))
)))
)

(smile-risorious
(definition (value (“to smile using the risorious muscles, which squares the mouth. A sign
of fear”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (tighten (theme (sem cheek)))))
        not (sem (cover (theme (sem (tooth)))))
)
)

(smile-contempt
(definition (value (“the mouth action of contempt, also called a sneer”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        tighten (theme (sem (lip)))
        raise (theme (sem (lip)))
)))
)

(smile-miserable
(definition (value (“a tense smile used to express suppressed negative emotion”)))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        tighten (theme (sem (lip)))
        pull-back (theme (sem (lip)))
)))
)
```

```
(blush
(definition (value ("the autonomic response that indicates blood is rushing to the head")))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        change-location (theme (sem (blood))) (destination (sem (face)))
        change-color (theme (sem (face)))
)))
)

(blanch
(definition (value ("the autonomic response that indicates blood has been sent away from
the head")))
(is-a (value (emotional-action)))
(has-event-as-part (sem (
        move-away (theme (sem (blood))) (source (sem (head)))
        whiten (theme (sem (face)))
)))
)
```

The two concepts blush and blanch provide an illustration of how inconsistently the ontology has been elaborated. The concept blanch benefited from two specialized concepts move-away and whiten, which helped in describing what precisely blanching is about. Blush on the other hand had to use the parent concepts of these two, change-location and change-color, because nothing more specific existed. I will not acquire anything more specific because I was able to sufficiently differentiate the two.

```
(laugh
 (IS-A (VALUE (emotional-action)))
(DEFINITION (VALUE ("to make the sounds and facial movements that express mirth,
ridicule, etc.")))
(has-event-as-part (sem (
        smile-zygomatic
        human-voice
        shed-tear
        breathe
)))
)
```

I have changed the parent concept of this legacy concept from EMOTIONAL-EVENT to emotional-action, thereby better situating it in the ontology. I have also removed the slot EXPERIENCER (SEM (HUMAN))) because this should be inherited from an ancestor concept.

(MOCK
(THEME (SEM (HUMAN)))
(IS-A (VALUE (opposition-event)))
(DEFINITION (VALUE ("to ridicule or to express it")))
)

This concept never had any business being in the emotional-event branch of the ontology. I have moved it out of my way, and removed two useless slots (EXPERIENCER (SEM (*NOTHING*))) and (AGENT (SEM (HUMAN))).

```
(humor
(definition (value (“to find something funny”)))
(is-a (value (emotional-event)))
(has-event-as-part (sem (
        laugh
        smile-zygomatic
        parasympathetic-response
)))
)
```

Here is the promised concept for humor, strangely—given Raskin’s research history—missing from the legacy ontology.

```
(SHED-TEAR
(IS-A (VALUE (emotional-action)))
(DEFINITION (VALUE ("to shed tears in response to something sad or joyful")))
(has-event-as-part (sem (
        produce (theme (sem water)) (source (sem (eye)))
        frown
        human-voice
        tighten (theme (sem (muscle (location (sem (chest esophagus))))))
)))
)
```

I have moved this legacy concept from EMOTIONAL-EVENT to emotional-action. I have also turned it into a small script, using what scanty concepts were available and appropriate. Note that the ontology is currently designed not to name many specific anatomic features such as the diaphragm and tear ducts, but instead to have an overarching concept such as MUSCLE that is specified by its location; in this case, the diaphragm is just a muscle in the chest. Tear ducts are not concepts in the ontology because practically, a person without trachoma is unaware that he or she has them except for the occasional empiric evidence of their existence, such as tearing of the eyes. Both of these work-arounds create an ontology that represents the world as understood by the average, mildly educated person, and also prevents the ontology from being littered with thousands of concepts that can be otherwise described. Computationally there is no difference between elaboration in the ontology and elaboration in the lexicon, but for the purpose of accommodating acquirers who are “sliding” around the ontology, it is better to keep concepts to a minimum.

(CONSOLE-EVENT
(IS-A (VALUE (abstract-social-activity)))
(DEFINITION (VALUE ("to make feel less sad or disappointed- to comfort")))
)

(INSULT
(IS-A (VALUE (abstract-social-activity)))
(DEFINITION (VALUE ("to subject to an act, remark, etc. meant to hurt the feelings or pride")))
)

(DEMORALIZE
(IS-A (VALUE (abstract-social-activity)))
(DEFINITION (VALUE ("to lower the morale or spirit of")))
)

(INTIMIDATE
(IS-A(VALUE (abstract-social-activity)))
(DEFINITION (VALUE ("to make afraid, as by threats")))
)

Though these legacy events happen most commonly when emotions are being experienced, I have moved them elsewhere in the ontology because they are not manifestations of emotion but rather reactions to them. I have also removed the slots (AGENT (SEM (HUMAN))), (BENEFICIARY (SEM (HUMAN))) and (EXPERIENCER (SEM (HUMAN))) when they occurred because these should, without doubt, be inherited from an ancestor concept.

```
(PITY
 (IS-A (VALUE (sadden)))
(DEFINITION (VALUE ("to feel for another's suffering or misfortune")))
(THEME (not (pity.agent)))
)
```

I have used the dot notation to great advantage in this concept; it allows me to specify that the theme of the sadness is anyone but the agent of the sadness. Under the branch of animates are animals and humans, which includes the social-role branch. To cover all the bases, I have added a second facet to this slot, allowing any object at all to be a possible object of pity. I have also made pity into a child concept of sadden. Therefore, the gist of the concept is that pity is a feeling of sadness for another—precisely what the original acquirer wrote in the definition.

This emotion among others requires the designation of a case role for the person feeling the emotion. It is an interesting question, discussed in chapter 5, of what case role exactly is appropriate to designate this person: AGENT would indicate that the person has control over what emotion he or she feels, which may or may not be true and in varying degrees, for instance with regard to socially conditioned emotions like nostalgia. Using the case role EXPERIENCER on the other hand would indicate that one may not choose to experience an emotion, but simply experiences it. This again would be appropriate if a true set of “native” or unconditioned emotions could be identified. More work for the psychologists.

```
(EXCITE
(SOURCE (SEM (EVENT)))
(IS-A (VALUE (EMOTIONAL-EVENT)))
(DEFINITION (VALUE ("to arouse the feelings of")))
(has-event-as-part (sem (sympathetic-response)))
)

(AGITATE
(PURPOSE (SEM (SHIFT-TREND)))
(IS-A (VALUE (POLITICAL-EVENT)))
(EXPERIENCER (SEM (HUMAN)))
(DEFINITION (VALUE ("to stir people up so as to produce changes")))
(has-event-as-part (sem (excite)))
)
```

I have removed the concept AGITATE from the emotional ontology entirely, seeing as its apparent purpose is to describe a political event, which parent it also had. I have also removed the slot (AGENT (SEM (HUMAN))) on the grounds that this should be inherited from an ancestor. From the concept EXCITE I have removed the SUBCLASSES slot because the removal of AGITATE left it empty, and also the (AGENT (SEM (ANIMAL))) slot because it should be inherited. I have added sympathetic-response as a component event for it.

```
(NEGATIVE-STATE
 (IS-A (VALUE (EMOTIONAL-EVENT)))
(DEFINITION (VALUE ("emotions, states or reactions generally regarded as negative,
such as scorn, hate, etc.")))
)
```

I am forced to leave NEGATIVE-STATE in the ontology, though I would remove it if I had the choice. Because I see no better place for it to go, I am leaving it under EMOTIONAL-EVENT, although it will probably be appropriate to move it elsewhere (or delete it). Really, I think this concept is best erased, especially given the concept SUFFER which will serve as a good parent for such concepts as pain, thirst, and hunger, which are negative states that provide motivation.

```
(POSITIVE-STATE
 (IS-A (VALUE (EMOTIONAL-EVENT)))
(DEFINITION (VALUE ("states, emotions or reactions generally regarded as positive,
such as happiness, love, etc.")))
)
```

As with NEGATIVE-STATE, this concept has been spared only because I am not allowed to delete it. Also as with NEGATIVE-STATE, I suspect it would be much more useful were motivation being acquired along with emotion, and would like to leave a note to future acquirers (possibly myself) to find something useful to do with these two.

```
(SUFFER
(IS-A (VALUE (passive-cognitive-event)))
(DEFINITION (VALUE ("to undergo pain, injury, etc.")))
)
```

I have moved the concept SUFFER out of the emotional-event branch because, though one can suffer from certain emotions, one can also suffer from any number of physical causes. This concept will become important if motivations are ever acquired, because suffering is principally caused by failure to satisfy motivations. I tried to add a has-event-as-part slot for this concept, but the ontology did not have concepts for pain, hunger, or thirst—so I decided that this concept stepped outside of my project's semantic box, and have left it.

```
(RESENT
(IS-A (VALUE (envy)))
(DEFINITION (VALUE ("to feel or to show hurt, bitterness, or indignation at")))
(theme (sem (object social-event living-event)))
)
```

I have moved RESENT from NEGATIVE-STATE to envy; it has a significantly smaller component of anticipation than does envy, and rather more anger. I deflect negative attention that may be caused by the placement of this concept from myself by reminding the reader that a concept's handle need not be semantically identical with the concept itself. I have narrowed the theme of this event to what I hope are a useful and truthful group of concepts; it is somewhat more specific than the ridiculous pair of fillers OBJECT and EVENT so often found in legacy concepts.

(HATE
(IS-A (VALUE (disgust anger)))
(DEFINITION (VALUE ("to have strong dislike or ill will for")))
(intensity (value (>.5)))
)

The difference between scorn and hate seems to be intensity. Were I building the ontology from scratch, I would not create two separate concepts, and in the lexicon would use the property intensity to differentiate the meanings of the two words. Since they are both already present, however, I will make them both children of disgust and anger, and differentiate with the intensity slot here in the ontology rather than in the lexicon. If a future acquirer is interested in exploring the use of both AGENT and EXPERIENCER, the two emotions might perhaps be differentiated this way—though in actual use the difference between the words is very much pragmatic, the handles chosen to represent concepts do not of course represent the essence of the concept perfectly.

(FRUSTRATE
(IS-A (VALUE (anger anticipation)))
(DEFINITION (VALUE ("to prevent from achieving a goal or satisfying a desire")))
(caused-by (sem (event)))
)

Attempting to define frustration in the context of Plutchik's theory and the legacy ontology brings out the serious ambiguity regarding what precisely lies between "anger" and "anticipation." Plutchik called it "aggressiveness," which I believe is a purposefully vague term. Already in the ontology we have dealt with ENVY and RESENT, both of which lie in this area; but then so surely does FRUSTRATION, which seems quite different from either of the others. Perhaps frustration can be thought of as a kind of resentment; but then when one envies one also resents the person who has what one wants. All three are quite closely tied up with motivation, perhaps more than any other branch—if someone would create a pretty pinwheel of motivations, the bridge between the two might originate right here. I have differentiated FRUSTRATE from ENVY and SCORN using slots.

(GRIEVE
(IS-A (VALUE (sadden anticipation remember)))
(DEFINITION (VALUE ("to mourn or suffer the loss of someone or something")))
(precondition (sem (lose)))
)

Plutchik himself calls grief the extreme end of the sadness continuum; we therefore have the option of differentiating SADDEN and GRIEVE using our favorite slot intensity. Grief, however, seems to have a special component of missing something that used to be and no longer is—it is tied up with the concepts of YEARN-FOR and FEEL-NOSTALGIC. I have, in this concept, defied Plutchik by blending two emotions that are not next to each other on his pinwheel. One of course is not anticipating something one is grieving for, but one wants it back—one would do something to regain it if one could. I therefore think my placement of the concept is appropriate. I have deleted two useless slots (THEME (SEM (EVENT OBJECT))) and (AGENT (SEM (*NOTHING*))).

```
(SHAME
(IS-A (VALUE (regret)))
(DEFINITION (VALUE ("to have a painful feeling of guilt for improper behavior,
etc.")))
(precondition (sem (event (caused-by (sem (shame.agent))))))
)
```

I am glad that the original acquirer included the word “guilt” in his definition of SHAME. I am not glad that this concept is basically identical to the concept REGRET, with the one difference that one can possibly regret something that cannot be blamed on oneself. I have differentiated shame from regret by indicating that its source is the person who experiences it.

Embarrassment would be represented by this concept, also, and it might be useful (later, during corpus testing) to create its own concept with slightly different dot notation to indicate that one can feel embarrassed/ashamed (again, don’t get hung up on the concept handles) about things that are not one’s own fault but that are closely connected to one, such as social status, family members, etc.

```
(WORRY
(IS-A (VALUE (anticipation have-fear)))
(DEFINITION (VALUE ("to make troubled or uneasy")))
(before (sem (event)))
)
```

This concept marks the second time that I have defied Plutchik by blending two emotions that are not next to each other on his wheel. Worry is clearly a mixture of anticipation and fear—so he has failed us in this. I have removed the useless (AGENT (SEM (*NOTHING*))) slot. I have tried to describe the temporal aspect of worry using the slot "before." In the same way that I argued regret can only be about the past, even if the thing in the past is a decision or certainty about the future, I argue that worry can only be about the future, even if the source of the worry is in the past.

Angst and the whole human condition could be scripted under this concept, and it would be a good project for a literary theory student to try doing just that based on his or her favorite theorist. Separate scripts for Kristeva, Foucault, etc. could be written and could in fact be extremely useful to the system if it was ever made to process fiction or cultural criticism texts—this assuming of course that the literary theory being worked with consisted of some valid insights into literature and the human condition.

(SYMPATHY
(DEFINITION (VALUE ("an affinity, association, or relationship between persons or things wherein whatever affects one similarly affects the other")))
(IS-A (VALUE (understand)))
(theme (sem (emotional-event)))
(beneficiary (not (sympathy.agent)))
)

I have moved this concept from POSITIVE-STATE to understand, a child of passive-cognitive-event, where the concept SUFFER also went. Sympathy is an understanding of emotions—so I have explained this in the slots. This is more than I owed the concept, given that it is no longer an emotion.

(FASCINATE
(IS-A (VALUE (anticipation surprise)))
(DEFINITION (VALUE ("to charm- captivate")))
)

Once again it is useful to blend two emotions that Plutchik did not associate with each other. I have also removed the useless slot (AGENT (SEM (ANIMAL))).

This is a troublesome concept. As will be discussed in chapter 5, it is a blend of two emotions which are diametrically opposite each other on Plutchik's wheel of emotions and which one therefore assumes would cancel each other out. How can a person both anticipate something and be surprised by it? The answer is that this concept is ripe for elaboration as a script, because just as the concept SURPRISE had a precondition of not knowing, fascination has a precondition of surprise and the result that one anticipates more of whatever surprised one to begin with it. Fascination is different from surprise in its temporal, enduring nature.

```
(APPRECIATE
(IS-A (VALUE (love)))
(DEFINITION (VALUE ("to show gratitude- to recognize gratefully")))
(intensity (value (=<.5)))
(has-event-as-part (sem (thank praise)))
)

(ENJOY
(IS-A (VALUE (passive-cognitive-event)))
(DEFINITION (VALUE ("to take pleasure in something or someone")))
)
```

Appreciation in the ontology labels the concept of gratefulness—thanks—admiration of a thing for its excellent qualities and the ways these qualities benefit oneself. I really cannot differentiate this concept from LOVE except in its intensity and, with some reservations, in a component of "thanking" though I see appreciation as aesthetic and moral also. I have removed the disastrously vague slot (THEME (SEM (EVENT OBJECT))) from APPRECIATE, and the useless slot (AGENT (SEM (*NOTHING*))) from ENJOY.

ENJOY itself will I think be more useful for the future acquirer of motivation than for myself, related as it is to reward systems. I see it as a twin concept of SUFFER (suffer being perhaps the evil twin), and so have moved it to the same place.

(EMPATHY
(DEFINITION (VALUE ("the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner")))
(IS-A (VALUE (involuntary-perceptual-event)))
(theme (sem (emotional-event)))
(beneficiary (not (empathy.agent)))
)

I have moved EMPATHY into the involuntary-perceptual-event branch of the ontology. Like SYMPATHY, it is not an emotion itself but rather a vicarious understanding of someone else's emotions. I have tried to represent the fine line between sympathy being a cognitive awareness of what another must feel, and empathy being actual vicarious experience.

The definition of this concept is a prime example of the work of an acquirer who doesn't understand that definitions are computationally useless. This person doubtless felt very proud of herself either for coming up with such a chunk of prose or for finding such an excellent definition in a dictionary. The net advantage to the system is, of course, nil. The concept had no new slots whatsoever and certainly not multiple inheritance to justify such an absence.

(COMPOSURE
(DEFINITION (VALUE ("a calmness or repose especially of mind, bearing, or appearance")))
(IS-A (VALUE (passive-cognitive-event)))
)

Plutchik calls the least intense end of the joy spectrum “serenity,” which is somewhat equivalent to composure. If I had to differentiate them I would say that serenity is the absence of fear, uncertainty, or disturbance of emotion, while composure is the absence of these things from one’s thoughts, or mind, or cognitive state. I have therefore moved COMPOSURE out of my own domain and into passive-cognitive-event. If/when a future acquirer becomes interested in religion, religious thought, alternative realities, transcendental meditation etc. this person may very well want to express most of that domain underneath a new concept called, for example, passive-cognitive-state (instead of event). I think it would be very tidy to say that an EMOTIONAL-EVENT may be experienced simply by virtue of being a member of the human race, while one might fiddle with one’s PASSIVE-COGNITIVE-STATE only by indulging in meditation or mind-altering chemicals.

(YEARN-FOR
(SUBCLASSES (VALUE (FEEL-NOSTALGIC)))
(IS-A (VALUE (desire)))
(DEFINITION (VALUE ("to be filled with longing for something that one does not necessarily have immediate access to")))
)

(FEEL-NOSTALGIC
(IS-A (VALUE (YEARN-FOR remember)))
(DEFINITION (VALUE ("to long for something that happened far away or long ago")))
)

It is too bad that the acquirer who created YEARN-FOR could not simply label it "want," because this is what the concept stands for. As I have made a point of differentiating emotion from motivation, I must move this concept (much as I like its label) out of my domain and into passive-cognitive-event, where there is a handy concept called DESIRE to be its foster parent. I have also removed the useless slot (THEME (SEM (OBJECT))) from this concept. Nostalgia, I think, is a combination of desire and remembrance, so I have added the second parent to its IS-A slot.

Whole books have been written about the history of nostalgia, and common knowledge among English majors is that it was manufactured sometime in the Victorian era or the Industrial Revolution (same time, different angles) either because this was the first time in human history that life changed rapidly enough for a person to miss "the old ways" within his or her own lifetime, or because publishing companies wanted to sell lots of sentimental novels (again, different angles on the same thing). Whether or not some emotions are socially "manufactured" or not would be an interesting topic for cogitation, though it is perhaps ultimately undeterminable.

Bookkeeping

Here begins the "bookkeeping" part of this ontological domain, made necessary because we have been working on paper and not in a concept editor: a list of concepts that have changed only in that they have gained or lost child concepts. The capital/lowercase distinction is not made here because, without exception, the only change made to these concepts has been in their SUBCLASSES slot.

(animal-living-event
(is-a (value (living-event)))
(definition (value ("physical-events which involve the functioning of an animal, including humans")))
(subclasses (value (wait stay sleep human-living-event breathe become-tired bleed excrete inhabit feeding-event offspring-event die autonomic-response)))
(agent (sem (animal))))

(emotional-event
(agent (sem (*nothing*)))
(definition (value ("emotional states or acts, such as anger, sadness, depression, happiness")))
(effect (sem (opposition-event mental-event emigrate criminal-activity entertain-event non-work-activity physical-event)))
(experiencer (sem (human)))
(is-a (value (mental-event)))
(opposite (inv (emotional-event)) (sem (emotional-event)))
(subclasses (value (happiness trust have-fear surprise sadden disgust anger anticipation humor excite negative-state positive-state)))
(domain-of (inv (experiencer))))

(abstract-social-activity
(agent (sem (human)))
(definition (value ("any social-event that shows no tangible result")))
(is-a (value (social-event)))
(subclasses (value (offend undertake satisfy resort-to refrain-from prevent popularize motivate intervene impress have-trust-in grow-accustomed-to facilitate deserve compensate cancel apply-for anticipate benefit-from commit-to deceive eliminate greet harass imitate indulge leave-work-activity open-to-public postpone reconcile register respect shun volunteer-event joke CONSOLE-EVENT INSULT DEMORALIZE INTIMIDATE))))

(external-vertebrate-part
(is-a (value (external-body-part)))
(definition (value ("external body part possessed by vertebrates but not by all vertebrates")))
(subclasses (value (nail-part face ear forehead skin eyebrow)))
(part-of-object (inv (turtle mammal bird amphibian vertebrate salamander fish reptile human))))

(opposition-event
(beneficiary (sem (human)))
(caused-by (inv (emotional-event)))
(definition
(value ("an event in which two or more parties face off in opposition or competition")))
(is-a (value (social-event)))
(subclasses (value (repudiate fight discriminate contradict compete avenge alienate abuse arm-event banish confront dare exploit interfere win MOCK))))

(political-event
(agent (inv (political-entity)) (default (governmental-role)) (sem (human)))
(definition (value ("An event that involves the governing body of a government or an organization")))
(is-a (value (social-event)))
(object-involved (inv (political-entity)))
(subclasses (value (vote torture tally-vote spy-on shift-trend revoke protest political-purge persecute nationalize liberate instate group-political-event emigrate demilitarize agitate decentralize designate government-activity industrialize labor-related-event lobby-for naturalize political-campaign poll rebel secede social-action subvert terrorist-activity veto run-for-office AGITATE)))
(timestamp (sem ("Modified Fri, Apr 11, 2003 by inna-Modified Wed, Apr 9, 2003 by inna"))))

(passive-cognitive-event
(agent (sem (*nothing*)))
(definition (value ("a cognitive action in which the cognater passively cognates")))
(domain-of (inv (experiencer)))
(experiencer (inv (animal)) (sem (human)))
(is-a (value (mental-event)))
(opposite (sem (*nothing*)))
(subclasses (value (desire understand know have-dream foresee confuse discover forget hope-for remember ignore prefer SUFFER ENJOY COMPOSURE)))
(timestamp (sem ("Modified Tue, Mar 18, 2003 by inna-Modified Tue, Mar 18, 2003 by inna"))))

(understand

(caused-by (sem (learn teach)))
(definition (value ("to comprehend the meaning and significance of something")))
(is-a (value (passive-cognitive-event)))
(subclasses (value SYMPATHY))))

(involuntary-perceptual-event (agent (sem (*nothing*)))
(definition (value ("a perceptual event in which the perceiver is not actively trying to perceive")))
(is-a (value (perceptual-event)))
(subclasses (value (involuntary-tactile-event involuntary-gustatory-event involuntary-auditory-event involuntary-olfactory-event involuntary-visual-event EMPATHY))))

(desire
(definition (value ("to want something")))
(is-a (value (passive-cognitive-event)))
(subclasses (value (YEARN-FOR)))
(timestamp (sem ("Modified Tue, Mar 18, 2003 by inna"))))

(remember
(definition (value ("To retain information in one's memory")))
(is-a (value (passive-cognitive-event)))
(theme (sem (object event)))
(subclasses (value (FEEL-NOSTALGIC GRIEVE))))

CHAPTER 5: DAMAGE CONTROL

Having looked over this ontology and its grain size, powers of description, and relative complexity, the reader may begin to appreciate how the ontology is in some ways extremely competent and in some ways woefully underdeveloped. In the early days of Raskin and Nirenburg's Mikrokosmos ontology, acquirers were given free reign to acquire as many lexical items as they could, at a low level of specification, using the bottom-up, i.e., from corpus to lexicon to ontology, methodology. The ontology therefore has several thousand solid and useful concepts which make acquisition of new concepts infinitely easier. For example, this exercise utilized concepts for many parts of the face and body that had already been acquired, concepts for causal relations and temporal ordering, and the occasional oddball like WRINKLE and SCREAM that added a nice touch to many of the concepts.

The ontology also has several areas which are finished with a high degree of polish. This always happens because a team has gotten a research grant to work on that particular area; as said before, past projects have included information security, tourism, medicine, finance, law, and sports. All domains require a large number of concepts that will be useful to other domains, so these specialized efforts always benefit the database as a whole. All domains also require a number of highly specialized concepts that will benefit no one else—an example is the concept SPORTS-GLOVE. This is a concept

sufficiently unlike any other that it may be differentiated (from other protective equipment used in sports, and from other kinds of gloves), and that also may have several child concepts. It is therefore a good idea to acquire this concept even though it is rather specific. On the other hand, some concepts really have no right to exist. SWIMMING-4X100M-MEDLEY-RELAY-MEN is such a one; there is no excuse to have not acquired this as a lexical item instead of as an ontological concept, and to have used more generic concepts to describe it accurately. However, it is in an obscure branch of the ontology and will not trip up an acquirer who is trying to find a concept by sliding down. Because it is so hidden and innocuous, it is not a candidate for deletion (as a bad concept would be if it was farther up in the ontology, where it would clutter up the already long lists of child concepts and make sliding a nightmare).

In many ways the ontology is also woefully inadequate. The human face has received a proportionately large amount of attention, not just by this acquirer, who worked in the medical domain, but by earlier acquirers working on whatever suited their fancy, because it is such a central object in the human mind. In spite of all this attention, there was no concept for "eyebrow," one of the most expressive parts of the face. We also had to acquire concepts for "raise" and "lower," because there were honestly no other concepts in the ontology that could be used to represent these two actions. There was no concept for "moving together" either, but the concept TIGHTEN worked well enough, so it was made to do. As one works in acquisition one is constantly running into these issues: one needs a concept that seems important and natural and unavoidable—and that doesn't exist in any way shape or form, even after dozens of people have done years of work on the database. Raskin and Nirenburg argue that knowledge database building is

not the unconquerable task many have thought it to be; whether this is true or not may be measured by the frequency of discovery of these holes in the knowledge base, as more acquisition work is done. The existence of these holes is also a powerful argument to put more resources towards building the core of concepts—resources that are rather harder to come by, unfortunately, than those for building up databases about a corporate sponsor's special interests. If nothing else, general database building is a good assignment for research assistants in NLP if they have been adequately trained to be trusted with the ontology..

Another way in which the ontology is surprisingly incomplete is the quality of description of many concepts, including those at quite high levels. In Chapter 4, I used lowercase letters to indicate what portions of concepts were newly acquired, leaving the legacy information in capital letters. Whenever I removed information from a concept, I noted what I had done. What the reader gleans from this is that the whole legacy EMOTIONAL-EVENT branch contained little or no semantic material: concepts had their IS-A and SUBCLASSES slots filled, a prosy definition, and possibly—in a few cases—a case-role or two filled with woefully vague over-inclusive fillers such as OBJECT and EVENT. Not only had no one tried to get at the essence of each concept, no one had even tried to differentiate them from each other. Please realize that in the new work of chapter 4, there are a few concepts which contain no slots of their own, but which inherit the slots of more than one parent concept. The unique combination of inherited slots makes these concepts themselves unique.

Of course, we have not gotten at what many would call the "essence" of each emotion, which is the component Nairne described as the "subjective experience" of it.

This is a hard lesson in ontological semantics, and one that a person must understand and make peace with: the semantic primitives of the ontology are the properties and *not* the subjective human understanding of the world. All the ontology really understands are things it can calculate (fillers that are numbers) and things that it can do statistical analysis of (all other fillers.) The system will never really know what the color red is to a human being, or what it is to be sad, to grieve, or to yearn. The computer is never going to be human, or conscious, or capable of perception on its own. We therefore must turn to the measurable and observable aspects of things, especially in such a very human domain as the emotions. Thus the concept elaboration focused on physiological and behavioral aspects of the emotions.

The legacy ontology proved invaluable in this project because it served as a catalogue of concepts that a large number of people over a large span of time have found to be significantly dissimilar from each other. As little effort as had been put into all of the emotions concepts, someone still made the effort to add each one to the database, and was therefore sufficiently motivated to do so by *something*. The acquirer therefore feels rather glad that she was not allowed to delete concepts at will, and raze the whole branch, as was her first instinct. In several instances the concepts were legitimate emotions that were not at all described by Plutchik's theory; in fact, we may have discovered that blending of Plutchik's primary emotions occurs not just between emotions which are next to each other on his pinwheel, but across any two given emotions. Leaving concepts intact (if not in place) also "flags" the branch and the concepts for future work by acquirers with a different perspective. Many, or most, of the people who have worked on the ontology are not from the North American Anglo/Germanic suburban white culture

that this particular acquirer comes from, and it is not just possible but highly likely that some day, someone will come along who more fully understands the distinctions that the concepts' original acquirers had in mind. For that acquirer, the concepts have been left standing.

The discovery that emotions may blend in ways Plutchik did not envision leads into possible future work in this domain. This project has been a preliminary one at best, creating an "ordinary" grain size ontology of new concepts come upon in the research that also accommodates the legacy ontology, which is inevitably and inherently less orderly than the research but also far more human and practical. What has been done here is really only the "top" part of top-down acquisition. In the future, the ontology must be tested against a corpus to be sure that it is competent. Without a doubt, it will initially fail this test, and more concepts will need to be added. As the ontology is tested the lexicon will be built, and this will be a far greater task—in the medical domain, lexical entries outnumber ontological entries ten to one, and it is reasonable to expect the difference to be even larger in the domain of emotion, given the English language's historic love of literature that depicts high emotional drama in excruciating detail.

Motivation will also have to be acquired, if emotion is to be truly complete. Many concepts formerly associated with emotion have been moved into other areas with the intention of incorporating them into motivation in the future. On days when one feels little faith in the human experience, one can argue that there is no emotion that *does not* stem from the satisfaction or frustration of a motivation, and even on one's best days one cannot argue that at least some instances of emotion *do*. If one is wet and cold and hungry

and alone, one surely cannot help feeling a little sad and afraid; if one is sharing a warm bed and a cup of tea, one cannot help but feel a little happy.

The question of what emotions are primitive—not just primary, but primitive—might be another avenue of research. Indeed, one could argue for a branch of "primitive" emotions that can be very easily linked with motivation, and for another branch of "cultivated" emotions that one is taught to experience. The acquirer thought of this often as she was working on chapter 4. Is not nostalgia supposed to be an emotion invented by the Victorians and carefully nurtured by the greeting card and knick-knack industries ever since? Is not romantic love supposed to have been invented by bored knight-errants shortly after they became quite certain there were no dragons to slay? What about appreciation, honor, shame, and embarrassment?

Of course there is no final answer to this problem. Freudians and Darwinists may argue themselves out of any corner and soundly tell one that yes indeed everything really does come down to sex (that's a motivation), while idealists and those of faith all over the world will argue that there are higher, nobler, selfless desires of the human heart that cannot be linked to base bodily needs and self-interest. One could then argue motivation from Maslow's position, that humans as social animals do indeed have more motivations that the Four Fs (fight, flight, feed, be fruitful) attributed to all animals, and that many emotions—including appreciation, honor, shame, and embarrassment—reflect these social motivations. The argument will never end. The ontologist is thankful for multiple inheritance, which allows everyone to be right.

To sum up, future work in this domain may include:

1) Refining the grainsize of existing concepts, in some cases by fleshing out their HAS-EVENT-AS-PART slots into full blown scripts
2) Testing the ontology against a corpus, which will result in
    a. Lexicon building
    b. More ontological concepts
    c. Justification of the previous work
3) Adding the domain of motivation, and grafting the two domains together
4) Diverse insights into why the legacy ontology was the way it was, leading back into number (1) and also (2) as lexicons for other languages are built
5) Elaboration of the ontology's structure, as new categories of emotions come into clearer focus.

In some cases, ontological organization of a domain is useful because, if one assumes that the domain should be orderly and somehow symmetrical, one can acquire what one believes to be the full domain and then see where there are empty spaces that ought to be filled with a concept. Plutchik's pinwheel model of emotion assumes precisely this—that the emotions smoothly blend into each other on one dimension, and become more or less intense in another. In Chapter 4, three emotions were discovered that are blends of emotions which are not next to each other on the wheel: grief is sadness and anticipation, worry is a blend of fear and anticipation, and fascination is a blend of anticipation and surprise. The first point is that, unless one understands what the concepts are in the context of the other concepts, the reason why these secondary emotions have been placed as they are is not intuitively obvious. The handles used for concepts are not fine-tuned semantic instruments; all the more reason for the ontologist to be vigilant. The

second point is that of the three blends of emotions listed above, two blend emotions which have two other emotions between them (grief and worry), and one actually blends emotions that are diametric opposites of each other (fascination). A superficial look at Plutchik's wheel gives one the idea that diametrically opposite emotions ought to cancel each other out. For example, one should not be able to be happy and sad at the same time. The example of fascination proves this wrong—when one is fascinated, one is astonished by what one sees and cannot wait to see more. One experiences both surprise and anticipation. Joy and sorrow are not separated, and are sometimes called "bittersweet." Anger and fear very often go together, as one hates the thing that threatens one. A blend of trust and disgust does not familiar to this acquirer, but one suspects—indeed expects—that some culture somewhere is familiar with it.

An acquirer looking for more trouble might therefore try a combination of each of Plutchik's eight primary emotions (and Plutchik has, of course, taken care of those emotions which are actually next door to each other). The acquirer would have to think carefully, though, about whether a particular blend of emotions really created a coherent emotional experience such as grief, worry, and fascination, or if the emotions really were two separate entities that could occur simultaneously in one human mind. Joy and sadness, for example, in a TMR would probably be given separate SOURCEs, even if the EXPERIENCER and BENEFICIARY are the same—a proud parent is happy that his adult offspring is moving on in life, and simultaneously sad that the baby is gone forever. Same experiencer, and conferred upon the same event, but for different reasons. Anger and fear might also be dually experienced in this way, rather than being true blends of each other. Indeed, the concept FASCINATE very well might need to be displaced by an

acquirer bent on making the ontology quite orderly with regard to Plutchik's theory. The current acquirer simply couldn't see how it was accommodated by the emotions named: fascination is not Plutchik's "awe" because it doesn't need to have a component of fear. Neither is it quite the same as "surprise" or even its more intense version "amazement." If one is fascinated one is simultaneously amazed by what one sees, and eager to see more. Thus the blend of diametrically opposed emotions.

Once again, just because my own Ameri-anglo-german worldview can't make two emotion concepts into one, it doesn't mean that no one can. I am just advocating that each acquirer do some soul-searching before creating Frankenstein sentiments. Indeed, at this point, when the legacy ontology has been combined with a theory of emotion that proposes itself to be complete, acquirers should probably not be looking for more emotional concepts to add but rather passively scouting for them as the infrastructure is tested against real life texts. Also, it is without a doubt a bad idea to assume that something so organic as emotion be held to a strict symmetrical standard. Indeed, an acceptable argument against Plutchik's theory is that it is simply too tidy to be true. Not only should we expect to find more emotions that are blends of what we already have, but we should also expect to find emotions that cannot be accommodated by mixing and matching—as was the case with several of the legacy concepts. Emotion ought to be a healthy organic system, and a healthy organic system sends out branches and feelers in all directions with no regard for aesthetic geometry, quite like a philodendron.

Given that the ontology is *somewhat* organized, however, it may be instructive to see how a few features have patterned within the emotions. If Plutchik's organization of emotions really does make them blend into one another, then it seems that the emotions

which caused sympathetic and parasympathetic responses ought to pattern together around his wheel, and it is for the most part so. HAPPINESS and TRUST both invoke a parasympathetic response, and are next to each other; the sympathetic response seems to pattern on either side of the parasympathetic, though, with ANTICIPATION and ANGER being both sympathetic as well as HAVE-FEAR and SURPRISE. Both SADDEN and DISGUST, diametrically opposite the two parasympathetic emotions, are without any particular autonomic response unless one considers the enteric response (nervous stomach) to be autonomic as many neuroscientists do, and indeed, being saddened and disgusted can both trigger it, although don't necessarily have to. The fact that the four sympathetic emotions do not pattern together is an interesting discovery.

The patterning of tears is extremely interesting, because they are associated with four primary emotions that are all separated from each other: one can cry from fear, sadness, anger, and happiness, but not from disgust, anticipation, trust, or surprise. These four emotions could of course lead to the other four: anticipation, as frustration, can turn to anger. Disgust frequently turns into fear. Trust might turn into happiness. Perhaps Plutchik's eight "primary" emotions could be profitably interpreted as four true primaries fear, sadness, anger, happiness, and four true secondary emotions disgust, anticipation, trust, and surprise—the "secondary" emotions such as love would then be true tertiary emotions. This possible re-interpretation of Plutchik's theory of emotion is the second interesting discovery. This discovery is in fact supported by the patterning of various kinds of smiles: one can smile with joy (the zygomatic smile), with anger (the risorious smile), with fear (also the risorious smile), and with sadness (the dampened or miserable smile). One does not smile from trust, surprise, or anticipation unless, as suggested

above, they lead into one of the four true primary emotions. There is however the problem of the contempt smile—the sneer—which patterns with disgust, not a true primary. Disgust is, in some cases, a natural reaction to a thing that is likely to cause harm in the form of disease or petty injury (one is disgusted by insects, by pus, by maggots, and by excrement). From this point of view, it could be called one of the motivations—and so could fear and anger, though probably not happiness (unless it is counted as an award for satisfying a motivation) or sadness, unless it results in motivation to move on (clean out one's closets, meet new people, etc) instead of simply enervating the experiencer. From an evolutionary standpoint, the advantages of weeping and melancholy are difficult to imagine.

Returning to the idea of a possible distinction between "native" and "social" emotions, one can make an argument to divide Plutchik's eight primaries along the same lines. Of the four true primaries, anger and fear seem to be the ones most likely to be experienced unintentionally, though it is certainly possible to "fan the flames" of someone's anger, or to make an intellectual effort to become even more afraid of something that one naturally would be. Perhaps manufactured fear could be called worry. Perhaps manufactured anger could be called hate. Happiness and sadness may be thrust upon one also, but because there is no strong autonomic response associated with them (happiness is parasympathetic, of course, but relaxation does not hit one in the same way adrenalin does) they are more dependent on conscious cooperation than anger and fear. Among the secondary emotions, disgust once again stands out as the one most likely to be experienced viscerally and unconsciously—anticipation and trust have significant conscious components, though surprise of course by its very nature can't be helped. The

third discovery is therefore that there does not appear to be any neat pattern, among Plutchik's eight primary emotions at least, of native and socialized emotions—all eight seem to have components of both.

LIST OF REFERENCES

Atallah, M. J., V. Raskin, M. Crogan, C. Hempelmann, F. Kerschbaum, D. Mohamed, and S. Naik 2001. Natural language watermarking: Design, analysis, and a proof-of-concept implementation. In: I. S. Moskowitz (ed.), **Information Hiding: 4th International Workshop, IH 2001, Pittsburgh, PA, USA, April 2001, Proceedings. Lecture Notes in Computer Science 2137**, Berlin: Springer, p. 185-199.

Atallah, M. J., V. Raskin, C. F. Hempelmann, M. Karahan, R. Sion, U. Topkara, and K. E. Triezenberg 2002. Natural language watermarking and tamperproofing. In: F. A. P. Petitcolas (ed.), **Information Hiding: 5th International Workshop, IH 2002, Noordwijkerhout, The Netherlands, October 2002, Revised Papers. Lecture Notes in Computer Science 2578**, Berlin: Springer, pp. 196-212.

Cacioppo, J. T., D. J. Klein, G. G. Berntson, and E. Hatfield 1993. The psychophysiology of emotion. In: M. Lewis and J. M. Haviland (eds.), **Handbook of emotions**. New York: Guilford Press, pp. 119-142.

Cannon, W. B. 1927. The James-Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology* 39, pp. 106-124.

Eibl-Eibesfeldt 1970. Ethology, the Biology of Behavior. New York: Holt, Rinehart and Winston.

Ekman, P. 1977. Biological and cultural contributions to body and facial movements. In: J. Blacking (ed.), **The Anthropology of the Body**, London: Academic Press.

Ekman, P. 1981. Methods for measuring facial action. In: K. Schere and P. Ekman (eds.), **Handbook on Methods of Nonverbal Communications Research**. New York: Cambridge University Press.

Ekman, P. 1992. **Telling Lies: Clues to Deceit in the Market Place, Politics, and Marriage, 2nd ed.**. New York: Norton.

Ekman, P., R. W. Levenson, and W. V. Frisen 1983. Autonomic nervous system activity distinguishes among emotions. *Science* 221, pp. 1208-1210.

Ellsworth, P. C. 1994. William James and emotion: Is a century of fame worth a century of misunderstanding*? Psychological Review* 101, 222-229.

Falk, C. 2005a. Lex Avoider: a lexicon display Windows-based tool. To be filed as a CERIAS Technical Report TR 2005-XXX.

Falk, C. 2005b. CAT: Concept Acquisition Tool. To be filed as a CERIAS Technical Report TR 2005-XXX.

Falk, C. 2005c. RAT: Lexicon Acquisition Tool. To be filed as a CERIAS Technical Report TR 2005-XXX.

Izard, C. E. 1994. Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin* 115, pp. 288-299.

James, W. 1890. The Principles of Psychology. New York: Holt. Reprinted: Cambridge, MA: Harvard University Press, 1983.

James, W. 1894. The physical basis of emotion. Psychological Review 1, pp. 516-529.

Kaada, B. 1967. Brain mechanisms related to aggressive behavior. In: C. D. Clemente and D. B. Lindsley (eds.), **Aggression and Defense**. Berkeley: University of California Press.

Kagan, J. 1997. Temperament and the reactions to unfamiliarity. *Child Development* 68, pp. 139-143.

Keltner, D., and B. N. Buswell 1996. Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition & Emotion* 10, pp. 155-171.

Krachina O. 2005. Role of ontological semantics in handling privacy policies. Poster presented at the 6th CERIAS Annual Research Symposium, Purdue University, West Lafayette, IN, March 23. TR 2005-87, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN

Lang, P. J. 1994. The varieties of emotional experience: A meditation on James-Lange theory. *Psychological Review* 101, pp. 211-221.

Levenson, R. W. 1992. Autonomic nervous system differences among emotions. *Psychological Science* 3, pp. 23-27.

Malaia, E. 2005. Digital identity management in ontological semantics: Methodology and Practice of Domain Representation. Poster presented at the 6th CERIAS Annual Research Symposium, Purdue University, West Lafayette, IN, March 23. TR

2005-88, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN

Maslow, A. H. 1954. **Motivation and Personality**. New York: Harper.

Matsumoto, D. 1987. The role of facial response in the experience of emotion: More methodological problems and a meta-analysis. *Journal of Personality and Social Psychology* 52, pp. 769774.

Nairne, J. S. 1997. **Psychology: The Adaptive Mind**. Pacific Grove, CA: Brooks/Cole.

Nirenburg, Sergei, and Victor Raskin 2004. **Ontological Semantics**. Cambridge, MA: MIT Press.

Oatley, K., and J. M. Jenkins 1996. **Understanding Emotions**. Cambridge, MA: Blackwell Publishers.

Ortony, A., and T. J. Turner 1990. What's basic about basic emotions? *Psychological Review* 97, pp. 315-331.

Oxford English Dictionary, www.oed.com, under site license to Purdue University.

Panksepp, J. 1982. Toward a general psychobiological theory of emotions. *Behavioral & Brain Sciences* 5, pp. 407-467.

Plutchik, R. (1994). **The Psychology and Biology of Emotion**. New York: Harper Collins College Publishers.

Raskin, V., C. H. Hempelmann, K. E. Triezenberg, and S. Nirenburg 2002. Ontology in Information Security: A Useful Theoretical Foundation and Methodological Tool. In: Victor Raskin and Christian F. Hempelmann (eds.), **Proceedings. New Security Paradigms Workshop 2001. 10-13 September 2001. Cloudcroft, NM**, New York: ACM Press, pp. 53-59.

Raskin, V., S. Nirenburg, C. F. Hempelmann, I. Nirenburg, and K. E. Triezenberg 2003. The genesis of a script for bankruptcy in ontological semantics. In: G. Hirst and S. Nirenburg (eds.), **Proceedings of the Workshop on Text Meaning, 2003 NAACL Human Language Technology Conference**, Edmonton, Canada, May 27-31.

Raskin, V., K. E. Triezenberg, O. Krachina, and E. Malaia 2005. Ontological semantic support for a specific domain. In: **Proceedings of the Second Annual Midwest Colloquium in Computational Linguistics**, Ohio State University, Columbus, OH. TR 2005-101, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN.

Russell, J.A. (1994*).* Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin* 115:1, pp. 102-141.

Schachter, S., and J. E. Singer 1962. Cognitive, social, and psysiological determinants of emotional state. *Psychological Review* 69, pp. 379-399.

Shaul, D. L., and N. L. Furbee 1998. **Language and Culture**. Prospect Heights, IL: Waveland Press.

Televnaja, J. 2004. Lexical Acquisition and Semantic Representation of English Phrasal Verbs in Ontological Semantics. An unpublished Ph.D. thesis, Program in Linguistics, Purdue University, W. Lafayette, IN.

Tomkins, S. S. 1962. **Affect, Imagery, and Consciousness, Vol. 1**. New York: Springer.

Triezenberg, K. E. 2003. From heuristics to ontological scripts. Workshop on Semantic Ontology-Ontological Semantics, University of Leipzig and DFG/IFOMIS, May 24-25. TR 2003-131, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN

Triezenberg, K. E. 2004. Ontology of the Medical Domain (500+ concepts). New York, NY: hakia, Inc.

Triezenberg, K. E. 2005. Ontological Semantic Lexicon for English Medical Domain (4,000 entries). New York, NY: hakia, Inc.

Triezenberg, K. E., and C. F. Hempelmann 2005. Methodology of Ontological Semantic Acquisition. A Tutorial Set for the Massive Ontological Semantic Acquisition Development, Moscow: Lomonosov University. TR 2005-112, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN

Appendix A

Lexicon Building Blocks

This file contains "building blocks" for ontological concepts and for noun, verb, and adjective lexical items. Whenever a line appears with double dashes
--like this—
it *must* be replaced with another building block. Whenever something inside a block appears in capital letters
LIKE THIS
It *must not* be changed. Whenever something appears in lowercase letters
like this
it *must* be changed to suit the specific item being acquired. Hopefully the items in lowercase letters are all sufficiently explanatory to allow this to be done easily.

These building blocks have been constructed to aid in the development of ontological and lexical items by hand, that is, without the KBAE software. The two principle difficulties in acquiring items by hand are, first, constructing the correct structure of each entry, and second, successfully nesting and matching parentheses. Copying and pasting the building blocks according to the double-dashed instructions will solve the first problem for most kinds of items that may be acquired. Carefully cutting and pasting each block in its entirety will solve the second problem.

Some of the building blocks are followed by notes on their expected usage.

For all blocks in this file, "category" may be filled in as follows:
V = verb
N = noun
ADJ = adjective

ONTOLOGICAL BUILDING BLOCKS:

An ontological concept:

(concept
(DEFINITION (VALUE (make up a definition)))
(IS-A (VALUE (insert the parent concept)))
(SUBCLASSES (VALUE (insert any subclasses the item may have; omit this slot if it has none)))
--insert other slots here--
)

An ontological slot:

(slot (facet (filler)))

A LEXICAL ITEM:

(lex-item
--insert senses here—
)

THE NOUN BUILDING BLOCKS

A Simple Noun Sense:

(lex-item (CAT N) (SYNONYMS "")
(ANNO (DEF"") (COMMENTS "") (EX ""))
(SYN-STRUC ((ROOT $VAR0) (CAT N)))
(SEM-STRUC (onto-concept
--insert slots here--
)))

A Slot:

(slot (facet filler))

A Slot for a Scalar Property:

(slot (VALUE value))

A Slot for a Binary Property:

(slot (VALUE value))

A Slot That References An Instantiated Event With Further Slots:

(slot (facet REFSEM1))

and after the SEM-STRUC,

(REFSEM1 (event (slot (facet filler))))

THE VERB BUILDING BLOCKS

A Verb Sense:

(verb-Vx (ANNO (COMMENT (indicate whether transitive or intransitive)))
(CAT V)
(SYN-STRUC(
(ROOT $VAR0) (CAT V)
--insert arguments here--
))
(SEM-STRUC (ont-concept
--insert slots here--
))
)

In the sem-structure for verbs, one expects to have a slot referencing each of the $var items that appear in the syn-struc. Most commonly, there is an AGENT slot that references the subject, a THEME slot that references the direct object, and a RELATION slot that references the object.

A Subject Argument:

(SUBJECT ((ROOT $VARx) (CAT N)))

A Direct Object Argument:

(DIRECTOBJECT ((ROOT $VARx) (CAT N)))

An Object Argument with Preposition:

(PP-ADJUNCT ((ROOT prep) (CAT PREP)
(OBJ ((ROOT root)))))

In object arguments with prepositions, there may be multiple ROOTs for both the PP-ADJUNCT and the OBJ. ROOT references a particular preposition for PP-ADJUNCT, and either a specific word or a $var for OBJ. Each ROOT appears inside its own pair of parentheses.

A Slot:

(slot (facet filler))

THE ADJECTIVE BUILDING BLOCKS

An Adjective Sense:

(lex-item-ADJx (CAT ADJ)
(ANNO (DEF "") (EX "") (COMMENTS ""))
(SYN-STRUC
(--insert ROOTs here--
))
(SEM-STRUC
--insert constraints here--
))

A Specific word ROOT:

(ROOT word) (CAT category)

A Variable Noun or Verb ROOT:

(ROOT $VARx) (CAT category)

A Variable Adjective ROOT:

(MODS ((ROOT $VARx) (CAT ADJ)))

The meaning of an adjective pertains to the meaning of the noun it modifies, so this MODS stuff in the attributive syn-struc captures: "about $var1, which is the head noun the adjective modifies"

A PP-adjunct ROOT:

(PP-ADJUNCT ((ROOT preposition) (CAT PREP) (OBJ ((ROOT $VARx)))))

A Comparative ROOT:

(SUBJECT (--insert subject ROOTs here--))
(DIRECTOBJECT (--insert direct object ROOTs here--))

A Slot Constraint:

(ont-concept
--insert slots here--
)

A Constraint That References a $var:

($VARx
--insert slots here--
)

Common slots to have here are INSTANCE-OF and DESCRIBES, which let you specify what a $var is supposed to be.

A Constraint That Describes a $var:

(^$VARx (attribute (value)))

A Slot:

(slot (facet filler))

ON RAPID PROPAGATION:

When acquiring lexical items, one often encounters groups of items that require similar or identical entries. In this situation, it is expedient to construct a template for one lexical entry, and then copy it for each lexical item, changing only the relevant fillers.

An example of a template constructed especially for a particular kind of lexical item: my "disease" template:

```
(
( (CAT N) (SYNONYMS “”)
(ANNO (DEF "") (COMMENTS "") (EX ""))
(SYN-STRUC ((ROOT $VAR0) (CAT N)))
(SEM-STRUC (
(caused-by (sem ))
(instrument-of (sem ))
(duration-typical (value ))
(has-symptom (sem ))
(treated-by (sem ))
(curability ())
(survivability ())
(treatability ())
(affect-quality-of-life (value ))
))))
```

Appendix B

emotional-event
 surprise
 sentiment
 pity
 laugh
  mock
 envy
 console-event
 excite
  agitate
 negative-state
  suffer
  scorn
  resent
  insult
  hate
  frustrate
  anger
  demoralize
  grieve
  have-fear
   intimidate
  regret
  sadden
  shame
  worry
 positive-state
  sympathy
  love
  fascinate
  appreciate
  enjoy
  happiness
  empathy
  composure
 shed-tear
 yearn-for
  feel-nostalgic

(EMOTIONAL-EVENT
  (DOMAIN-OF
    (INV (EXPERIENCER)))

```
  (SUBCLASSES
    (VALUE (CONSOLE-EVENT ENVY EXCITE LAUGH NEGATIVE-STATE PITY
POSITIVE-STATE SENTIMENT SHED-TEAR SURPRISE YEARN-FOR)))
  (OPPOSITE
    (SEM (EMOTIONAL-EVENT))    (INV (EMOTIONAL-EVENT)))
  (IS-A
    (VALUE (MENTAL-EVENT)))
  (EXPERIENCER
    (SEM (HUMAN)))
  (EFFECT
    (SEM (CRIMINAL-ACTIVITY EMIGRATE ENTERTAIN-EVENT MENTAL-
EVENT NON-WORK-ACTIVITY OPPOSITION-EVENT PHYSICAL-EVENT)))
  (DEFINITION
    (VALUE ("emotional states or acts, such as anger, sadness, depression, happiness")))
  (AGENT
    (SEM (*NOTHING*)))
)

(SURPRISE
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to come upon suddenly or unexpectedly- to take unawares")))
  (AGENT
    (SEM (ANIMAL)))
)

(SENTIMENT
  (IS-A
    (VALUE (EMOTIONAL-EVENT OPINION)))
  (DEFINITION
    (VALUE ("a complex mix of feelings and opinions")))
)

(PITY
  (THEME
    (SEM (HUMAN)))
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to feel for another's suffering or misfortune")))
)

(LAUGH
  (SUBCLASSES
```

```
    (VALUE (MOCK)))
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (EXPERIENCER
    (SEM (HUMAN)))
  (DEFINITION
    (VALUE ("to make the sounds and facial movements that express mirth, ridicule,
etc.")))
)

(MOCK
  (THEME
    (SEM (HUMAN)))
  (IS-A
    (VALUE (LAUGH)))
  (EXPERIENCER
    (SEM (*NOTHING*)))
  (DEFINITION
    (VALUE ("to ridicule or to express it")))
  (AGENT
    (SEM (HUMAN)))
)

(ENVY
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to feel discontent over or desire for other people's possessions,
accomplishments, etc.")))
)

(CONSOLE-EVENT
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to make feel less sad or disappointed- to comfort")))
  (AGENT
    (SEM (HUMAN)))
)

(EXCITE
  (SUBCLASSES
    (VALUE (AGITATE)))
  (SOURCE
    (SEM (EVENT)))
```

```
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to arouse the feelings of")))
  (AGENT
    (SEM (ANIMAL)))
)

(AGITATE
  (PURPOSE
    (SEM (SHIFT-TREND)))
  (IS-A
    (VALUE (EXCITE POLITICAL-EVENT)))
  (EXPERIENCER
    (SEM (HUMAN)))
  (DEFINITION
    (VALUE ("to stir people up so as to produce changes")))
  (AGENT
    (SEM (HUMAN)))
)

(NEGATIVE-STATE
  (SUBCLASSES
    (VALUE (ANGER DEMORALIZE FRUSTRATE GRIEVE HATE HAVE-FEAR
INSULT REGRET RESENT SADDEN SCORN SHAME SUFFER WORRY)))
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("emotions, states or reactions generally regarded as negative, such as scorn,
hate, etc.")))
)

(SUFFER
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to undergo pain, injury, etc.")))
)

(SCORN
  (THEME
    (SEM (OBJECT)))
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
```

```
    (VALUE ("to regard or refuse with extreme contempt")))
)

(RESENT
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to feel or to show hurt, bitterness, or indignation at")))
)

(INSULT
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (EXPERIENCER
    (SEM (*NOTHING*)))
  (DEFINITION
    (VALUE ("to subject to an act, remark, etc. meant to hurt the feelings or pride")))
  (BENEFICIARY
    (SEM (HUMAN)))
  (AGENT
    (SEM (HUMAN)))
)

(HATE
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to have strong dislike or ill will for")))
)

(FRUSTRATE
  (SOURCE
    (SEM (EVENT)))
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to prevent from achieving a goal or satisfying a desire")))
  (AGENT
    (SEM (HUMAN)))
)

(ANGER
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
```

```
    (VALUE ("to create or have hostile feelings because of opposition, hurt, etc.")))
)

(DEMORALIZE
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to lower the morale or spirit of")))
)

(GRIEVE
  (THEME
    (SEM (EVENT OBJECT)))
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to mourn or suffer the loss of someone or something")))
  (AGENT
    (SEM (*NOTHING*)))
)

(HAVE-FEAR
  (SUBCLASSES
    (VALUE (INTIMIDATE)))
  (IS-A
    (VALUE (NEGATIVE-STATE)))
  (DEFINITION
    (VALUE ("to feel afraid")))
  (AGENT
    (SEM (*NOTHING*)))
)

(INTIMIDATE
  (IS-A
    (VALUE (HAVE-FEAR)))
  (EXPERIENCER
    (SEM (HUMAN)))
  (DEFINITION
    (VALUE ("to make afraid, as by threats")))
  (AGENT
    (SEM (HUMAN)))
)

(REGRET
  (THEME
```

```
   (SEM (EVENT)))
  (IS-A
   (VALUE (NEGATIVE-STATE)))
  (DEFINITION
   (VALUE ("to feel sorry about (an event, one's acts, etc.)")))
)

(SADDEN
  (IS-A
   (VALUE (NEGATIVE-STATE)))
  (DEFINITION
   (VALUE ("to make or become sad")))
)

(SHAME
  (IS-A
   (VALUE (NEGATIVE-STATE)))
  (DEFINITION
   (VALUE ("to have a painful feeling of guilt for improper behavior, etc.")))
)

(WORRY
  (IS-A
   (VALUE (NEGATIVE-STATE)))
  (DEFINITION
   (VALUE ("to make troubled or uneasy")))
  (AGENT
   (SEM (*NOTHING*)))
)

(POSITIVE-STATE
  (SUBCLASSES
   (VALUE (APPRECIATE COMPOSURE EMPATHY ENJOY FASCINATE
HAPPINESS LOVE SYMPATHY)))
  (IS-A
   (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
   (VALUE ("states, emotions or reactions generally regarded as positive, such as
happiness, love, etc.")))
)

(SYMPATHY
  (DEFINITION
   (VALUE ("an affinity, association, or relationship between persons or things wherein
whatever affects one similarly affects the other")))
```

```
  (IS-A
    (VALUE (POSITIVE-STATE)))
)

(LOVE
  (THEME
    (SEM (OBJECT)))
  (IS-A
    (VALUE (POSITIVE-STATE)))
  (DEFINITION
    (VALUE ("to feel strong affection for")))
)

(FASCINATE
  (IS-A
    (VALUE (POSITIVE-STATE)))
  (DEFINITION
    (VALUE ("to charm- captivate")))
  (AGENT
    (SEM (ANIMAL)))
)

(APPRECIATE
  (THEME
    (SEM (EVENT OBJECT)))
  (IS-A
    (VALUE (POSITIVE-STATE)))
  (DEFINITION
    (VALUE ("to show gratitude- to recognize gratefully")))
)

(ENJOY
  (IS-A
    (VALUE (POSITIVE-STATE)))
  (DEFINITION
    (VALUE ("to take pleasure in something or someone")))
  (AGENT
    (SEM (*NOTHING*)))
)

(HAPPINESS
  (IS-A
    (VALUE (POSITIVE-STATE)))
  (DEFINITION
    (VALUE ("having, showing, or causing great pleasure or joy")))
```

```
)

(EMPATHY
  (DEFINITION
    (VALUE ("the action of understanding, being aware of, being sensitive to, and
vicariously experiencing the feelings, thoughts, and experience of another of either the
past or present without having the feelings, thoughts, and experience fully communicated
in an objectively explicit manner")))
  (IS-A
    (VALUE (POSITIVE-STATE)))
)

(COMPOSURE
  (DEFINITION
    (VALUE ("a calmness or repose especially of mind, bearing, or appearance")))
  (IS-A
    (VALUE (POSITIVE-STATE)))
)

(SHED-TEAR
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to shed tears in response to something sad or joyful")))
)

(YEARN-FOR
  (THEME
    (SEM (OBJECT)))
  (SUBCLASSES
    (VALUE (FEEL-NOSTALGIC)))
  (IS-A
    (VALUE (EMOTIONAL-EVENT)))
  (DEFINITION
    (VALUE ("to be filled with longing for something that one does not necessarily have
immediate access to")))
)

(FEEL-NOSTALGIC
  (IS-A
    (VALUE (YEARN-FOR)))
  (DEFINITION
    (VALUE ("to long for something that happened far away or long ago")))
)
```

VITA

Katrina Elise Triezenberg
441 Heavilon Hall
Purdue University
500 Oval Drive
West Lafayette, Indiana 47907-2038
e-mail: kattriez@purdue.edu
http://expert.ics.purdue.edu/~kattriez/cv.pdf

Education

2005: Ph.D. Dissertation "Ontology of Emotion," supervised by Victor Raskin, defended on September 24

2005: Ph.D. Candidate in Linguistics, Purdue University

2004: M.A. in Linguistics, Purdue University
Thesis, "New Philology: A Call For (Inter)Disciplinary Action," supervised by Victor Raskin

2004: Certificate in Natural Language Processing, Purdue University

2003: B.S. in Biology, Minor in Psychology, Purdue University

2003: B.A. in English, Purdue University

1998: Honors degree, Homestead High School, Fort Wayne, Indiana

Employment

2005-6 Supervisor, Massive Ontological Semantic Acquisition Development, Hakia, Inc., New York, NY, and Moscow, Russia

2004- Senior Consultant in Ontological Semantics, hakia, Inc., New York, NY

2004-5 Teaching Assistant, Department of English, Purdue University.

2005 Co-Director and Senior Presenter, Lexical Acquisition School, Moscow State Lomonosov University, Moscow, Russia

2001 - 2003: Research Assistant, Natural Language Processing Laboratory, Interdepartmental Program in Linguistics, and CERIAS, Purdue University.

2000: Renewal service operator, Purdue University Libraries, West Lafayette, Indiana.

Languages

English - native
French - working knowledge
Italian - working knowledge
Swahili - linguistic familiarity

Classes Taught

Fall 2005: Graduate Instructor, ENGL 227/LING 201: Elements of Linguistics, Purdue University

Spring 2005: Graduate Instructor, ENGL 227/LING 201: Elements of Linguistics, Purdue University

Spring 2005: Co-Instructor, LING 689O: Ontological Acquisition

Fall 2004: Teaching Assistant, ENGL 506: Introduction to Linguistics, Purdue University

Fall 2003: Guest lecturer, CS 590E: Topical Lectures in Information Security (Prof. Ninghui Li), Purdue University

Fall 2003: Supplemental instructor, LING 689N/ENGL 628: Natural Language Processing (Prof. Victor Raskin), Purdue University

Spring 2003: Substitute instructor, ENGL 227/LING 201: Elements of Linguistics (Prof. Victor Raskin), Purdue University

2002-2003: Supplemental instructor and Group Leader (with Christian F. Hempelmann), ENGL/LING 590: Ontological Semantics (for Prof. Victor Raskin's NLP Research Group; 8 graduate and undergraduate students of ELL/ENGL and LING)

Fall 2002: Substitute instructor, ENGL 506/LING 500: Introduction to General and English Linguistics (Prof. Victor Raskin), Purdue University

Spring 2002: Substitute instructor, ENGL 511/LING 531: Semantics (Prof. Victor Raskin), Purdue University

Fall 2001: Supplemental instructor with Christian Hempelmann, ENGL 628/LING 689N: Natural Language Processing (Prof. Victor Raskin), Purdue University

Publications

Ontological semantic support for a specific domain. In: **Proceedings of the Second Annual Midwest Colloquium in Computational Linguistics**, Ohio State University, Columbus, OH. TR 2005-112, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN (with V. Raskin, E. Malaia, and O. Krachina).

Methodology of Ontological Semantic Acquisition. A Tutorial Set for the Massive Ontological Semantic Acquisition Development, Moscow: Lomonosov University, 2005. TR 2005-101, a Technical Report, Center for Education and Research in Information Assurance and Security (CERIAS), Purdue University, West Lafayette, IN (with Christian F. Hempelmann).

Re-Examining the GTVH. *Humor* 17:4, 2004.

The Purdue Ontological Semantic Project. In: Damir Cavar and Paul Rodriguez (eds.), **Proceedings of the First Annual Midwest Colloquium in Computational Linguistics, Indiana University**, Bloomington, IN, June 25-26, 2004 (with Victor Raskin, Krista Bennett, Christian F. Hempelmann, Evguenia Malaya, and Dina Mohamed)

A Linguistic Theory Will Not Suffice for Literary Humor. In: **Abstracts of ISHS 2004: Annual Meeting of the International Society for Humor Studies**. Instituts Universitaires de Technologie at Dijon, France, June.

Getting Sophisticated about Sophistication: Inference at the Service of Humor. In: **Abstracts of ISHS 2003: Annual Meeting of the International Society for Humor Studies**. Chicago: Northeastern Illinois University, July 2003. With Victor Raskin.

The Genesis of a Script for Bankruptcy in Ontological Semantics. In: G. Hirst and S. Nirenburg (eds.), **Proceedings of the Workshop on Text Meaning, 2003 NAACL Human Language Technology Conference**, Edmonton, Canada, May 27-31. With Victor Raskin, Sergei Nirenburg, Inna Nirenburg, and Christian F. Hempelmann.

Levels of Sophistication in Humor Intelligence Agents. In: A. Nijholt (ed.), **Proceedings of the Humor Interface Workshop at CHI-2003: Computer Humor Interface Conference. Fort Lauderdale, FL, April 6, 2003**. Einschede: University of Twente. With Victor Raskin.

Natural Language Watermarking and Tamperproofing. In: Fabien A. P. Petitcolas (ed.), **Information Hiding, 5th International Workshop, IH 2002**, Berlin: Springer, 2002, pp. 196-210. With Mikhail J. Atallah, Victor Raskin, Christian F. Hempelmann, Mercan Karahan, Radu Sion, and Umut Topkara.

Why NLP should move into IAS. In: Steven Krauwer (ed.), **Proceedings of the Workshop on a Roadmap for Computational Linguistics**, Taipei, Taiwan: Academia Sinica, 2002, pp. 1-7. With Victor Raskin, Sergei Nirenburg, Mikhail J. Atallah, and Christian F. Hempelmann.

Ontology in Information Security: A Useful Theoretical Foundation and Methodological Tool. In: Victor Raskin and Christian F. Hempelmann (eds.), **Proceedings. New Security Paradigms Workshop 2001. 10-13 September 2001. Cloudcroft, NM**, New York: ACM Press, 2002 pp. 53-59. With Victor Raskin, Christian F. Hempelmann, and Sergei Nirenburg.

Presentations

Ontological Semantics of Humor. A Pre-Conference Tutorial. ISHS-05: International Conference on Humor Research, Youngstown State University, Youngstown, OH, 2005

Ontological Semantic Support for a Specific Domain. MCLC-05: The Second Annual Midwest Colloquium on Computational Linguistics, Ohio State University, Columbus, OH, 2005 (with Victor Raskin, Olga Krachina, and Evgueniya Malaia).

A Linguistic Theory Will Not Suffice for Literary Humor. Linguistics Lunches Series, Linguistics Program, Purdue University, West Lafayette, IN, October 13, 2004.

The Purdue Ontological Semantic Project. The First Annual Midwest Colloquium in Computational Linguistics, Indiana University, Bloomington, IN, June 25-26, 2004 (with Victor Raskin, Krista Bennett, Christian F. Hempelmann, Evguenia Malaia, and Dina Mohamed)

A Linguistic Theory Will Not Suffice for Literary Humor. ISHS 2004: International Conference on Humor Research, Instituts Universitaires de Technologie at Dijon, France, June.

New Philology: A Call For (Inter)Disciplinary Action. Linguistics Lunches Series, Linguistics Program, Purdue University, West Lafayette, IN, March 10, 2004.

Getting Sophisticated About Sophistication: Inference at the Service of Humor. Keynote address, ISHS 2003: International Conference on Humor Research, NEIU at Chicago, Illinois, USA, July. With Victor Raskin.

The Genesis of a Script for Bankruptcy in Ontological Semantics. Workshop on Text Meaning, 2003 NAACL Human Language Technology Conference, Edmonton, Canada, May 27-31. With Victor Raskin, Sergei Nirenburg, Inna Nirenburg, and Christian F. Hempelmann.

From Heuristics to Ontological Scripts. Workshop on Semantic Ontology-Ontological Semantics, University of Leipzig and DFG/IFOMIS, May 24-25, 2003.

Natural Language Information Assurance and Security. Fourth Annual CERIAS Research Symposium, West Lafayette, IN, April 8-9 2003. With Mikhail J. Atallah, Victor Raskin, Christian F. Hempelmann, Krista Bennett, Mercan Karahan, and Umut Topkara.

Ontological Terminology Standardization. Poster session. Fourth Annual CERIAS Research Symposium, West Lafayette, IN, April 8-9 2003.

Semantic Mimicking. Poster session. Fourth Annual CERIAS Research Symposium, West Lafayette, IN, April 8-9 2003.

Levels of Sophistication of Humor Intelligent Agents. Workshop on Humor Modeling in the Interface, CHI-2003, Fort Lauderdale, FL, April 5-10 2003. With Victor Raskin.

Natural Language Watermarking and Tamperproofing. Third Annual CERIAS Research Symposium, West Lafayette, IN, April 16-17 2002. With Mikhail J. Atallah, Victor Raskin, Christian F. Hempelmann, Mercan Karahan, Radu Sion, and Umut Topkara.

"Ontology for IAS: Extending Research and Application Paradigms in Information Security." Invited presentation at the Center for Education and Research in Information Assurance and Security (CERIAS) Security Seminar, Purdue University, February 6 2002. With Victor Raskin and Christian F. Hempelmann. Web re-broadcast on several occasions.

Online Resources

2005 Ontology for the Domain of Emotion (50+ concepts)

2005 Ontological Semantic Lexicon for English Medical Domain (4,000 entries)

2004 Ontology for the Medical Domain (500+ concepts)

2002 Ontology for the Animal Branch (30+ concepts), Computing Research Laboratory, New Mexico State University, May

2001 Ontology and English Lexicon for the Domain of Information Security (with Christian F. Hempelmann)

Memberships

2005- Linguistic Society of America

2002 - 2003: Vice President of the Purdue Linguistics Association <http://expert.ics.purdue.edu/~linguist>

Recognition and Scores

2004-2006: Purdue University Doctoral Fellowship

2004-2005: Lynn Fellowship (declined)

2004: M.A. in Linguistics GPA: 4.0/4.0

2003-2004: Purdue University Special Initiatives Fellowship for Graduate Studies in Linguistics

2003: Graduating Seniors' Certificate of Achievement, Department of English, Purdue University, May.

2003: Purdue Literary Awards
1st Place Kneale Award for Linguistics
1st Place Kneale Award for Cultural Criticism
Honorable Mention, Helen Bass Williams Award

2003: Cumulative Bachelor's GPA: 3.62/4.0, B.A. GPA: 4.0/4.0

2003: At graduation, Semester Honors seven times, Dean's List eight times

2003: Research Internship, ITC-irst: Institute for Researcfh in Science and Technology, Trento, Italy. Supervisor: Dr. Oliviero Stock

2002: GRE-CAT Verbal: 700, Quantitative: 750, Analytical: 5.0

2002: Undergraduate internship at New Mexico State University's

Computing Resource Laboratory. Supervisors, Drs. Sergei Nirenburg and Marjorie McShane

Volunteer Work

2001: Website design and execution, International Awareness Week: Celebrating World Cultures, for Purdue University.

2001: Substitute presenter for the World Cinema Series, Department of History, Purdue University.

2000 - 2001: Undergraduate research assistant, Robinson Laboratory, Department of Biology, Purdue University.

2000 - 2001: Website design and execution, The International Center of Greater Lafayette, Inc.