NORTHWESTERN
UNIVERSITY

Electrical Engineering and Computer Science Department

Technical Report
NWU-EECS-09-17
July 30, 2009

A Pragmatic Approach to Computational Narrative Understanding

Emmett Tomai

Abstract

Narrative understanding is a hard problem for artificial intelligence that requires deep semantic understanding of natural language and broad world knowledge. Early research in this area stalled due to the difficulty of knowledge engineering and a trend in the field towards robustness at the expense of depth. This work explores how a practical integration of more recent resources and theories for natural language understanding can perform deep semantic interpretation of narratives when guided by specific pragmatic constraints. It shows how *cognitive models* can provide pragmatic context for narrative understanding in terms of well-defined reasoning tasks, and how those tasks can be used to guide interpretation and evaluate understanding. This work presents an implemented system, EA NLU, which has been used to interpret narrative text input to cognitive modeling simulations. EA NLU integrates existing large-scale knowledge resources with a controlled grammar and a compositional semantic interpretation process to generate highly expressive logical representations of sentences. Delayed disambiguation and representations from dynamic logic are used to separate this compositional process from a query-driven discourse interpretation process that is guided by pragmatic concerns and uses world knowledge. By isolating explicit points of ambiguity and using *limited evidential abduction*, this query-driven process can automatically identify the disambiguation choices that entail relevant interpretations. This work shows how this approach maintains computational tractability without sacrificing expressive power. EA NLU is evaluated through a series of experiments with two cognitive models, showing that it is capable of meeting the deep reasoning requirements those models pose, and that the constraints provided by the models can effectively guide the interpretation process. By enforcing consistent interpretation principles, EA NLU benefits the cognitive modeling experiments by reducing the opportunities for tailoring the input. This work also explores the use of a theory of *narrative functions* as a heuristic guide to interpretation in EA NLU. In contrast to potentially global task-specific queries, these narrative functions can be inferred on a sentence-by-sentence basis, providing incremental disambiguation. This method is evaluated by interpreting a set of Aesop's fables, and showing that the interpretations are sufficient to capture the intended lesson of each fable.

Keywords: Pragmatics, natural language understanding, knowledge representation, abduction, narrative

NORTHWESTERN UNIVERSITY

A Pragmatic Approach to Computational Narrative Understanding

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Emmett Russell Tomai

EVANSTON, ILLINOIS

December 2009

ABSTRACT

A Pragmatic Approach to Computational Narrative Understanding

Emmett Russell Tomai

Narrative understanding is a hard problem for artificial intelligence that requires deep semantic understanding of natural language and broad world knowledge. Early research in this area stalled due to the difficulty of knowledge engineering and a trend in the field towards robustness at the expense of depth. This work explores how a practical integration of more recent resources and theories for natural language understanding can perform deep semantic interpretation of narratives when guided by specific pragmatic constraints. It shows how *cognitive models* can provide pragmatic context for narrative understanding in terms of well-defined reasoning tasks, and how those tasks can be used to guide interpretation and evaluate understanding.

This work presents an implemented system, EA NLU, which has been used to interpret narrative text input to cognitive modeling simulations. EA NLU integrates existing large-scale knowledge resources with a controlled grammar and a compositional semantic interpretation process to generate highly expressive logical representations of sentences. Delayed disambiguation and representations from dynamic logic are used to separate this compositional process from a query-driven discourse interpretation process that is guided by pragmatic concerns and uses world knowledge. By isolating explicit points of ambiguity and using *limited evidential abduction*, this query-driven process can automatically identify the disambiguation choices that entail relevant interpretations. This work shows how this approach maintains computational tractability without sacrificing expressive power. EA NLU is evaluated through a series of experiments with two cognitive models, showing that it is capable of meeting the deep reasoning requirements those

models pose, and that the constraints provided by the models can effectively guide the interpretation process. By enforcing consistent interpretation principles, EA NLU benefits the cognitive modeling experiments by reducing the opportunities for tailoring the input.

This work also explores the use of a theory of *narrative functions* as a heuristic guide to interpretation in EA NLU. In contrast to potentially global task-specific queries, these narrative functions can be inferred on a sentence-by-sentence basis, providing incremental disambiguation. This method is evaluated by interpreting a set of Aesop's fables, and showing that the interpretations are sufficient to capture the intended lesson of each fable.

ACKNOWLEDGEMENTS

I would like to thank Ken Forbus for advising me in this work. For sticking with me and supporting me through several different research directions and helping me to find my way through to completion. Thanks also to my committee members, Chris Riesbeck and Ian Horswill, who were always willing to take the time to sift through my ideas and problems and provide insightful feedback.

I am also grateful to the members of the Qualitative Reasoning Group with whom I have shared offices, meetings and work over the past seven years. I am particularly grateful to Kate Lockwood, Morteza Dehghani and Matt Klenk. Our dedicated support group, committed to seeing us all graduate and find jobs, was not only crucial to my finishing, but one of the highlights of my graduate school experience.

There are too many close friends who have been our family here in Chicago to name them all, but a heartfelt thanks to everyone who gave me a life outside of graduate school to keep some perspective. I want to specially mention our small group and friends at Good News Bible Church as well as Eddie Yoon, Dan Voelker and Sam Cheng. You guys were there for the highs and the lows and I hope you all know how much I appreciated your company along the way.

To my parents, Elbert and Janet Tomai, it would be silly for me to try and say in a couple sentences all that we've talked about in the past two and half years. I am glad that this experience helped me see just how precious it is to have parents that still have advice, support and encouragement for me any time I call. To my sister and brother-in-law, Jennifer and Daniel Moy, I am thankful that we have been able to continue sharing life despite the distance. I take

lots of joy, encouragement and inspiration from seeing how you raise your kids, wrestle with decisions and deal with life.

To my brother, Elliott Tomai, who struggled far longer and harder than I ever did. Our light and momentary troubles may be passing, but I would have liked to have walked out of that darkness together. I would much rather be at your graduation than mine. Thank you for sharing the journey and for being my brother. You are missed.

To Evan and Corrie, you'll barely remember these years, but when I think back on Chicago and graduate school the first thing I'll remember is you. I love you guys and I can't wait to be working close to home again. To my wife, Deborah, the best thing that ever happened to me. You went way beyond being understanding, you gave me the energy and purpose to press on day by day even while you were heroically caring for our kids and building a life for us. I can't thank you enough for loving me so well. We made it!

From the beginning to the end, I thank God for the grace by which I stand. I have no other hope than this, that Jesus loves me and he is faithful.

But he said to me, "My grace is sufficient for you, for my power is made perfect in weakness." Therefore I will boast all the more gladly of my weaknesses, so that the power of Christ may rest upon me.

Table of Contents

List of Figures

List of tables

1. Introduction

Narrative is a fundamental form of human linguistic communication, where the narrator describes a series of events in such a way that draws attention to the relationships between them. The hearer is expected to infer a great deal more information about the story than is contained in the explicit details provided. Early work in narrative understanding in the field of artificial intelligence demonstrated the crucial role of world knowledge in this inferential process. Charniak's model of answering questions about children's stories (Charniak, 1972) assumed that the interesting questions were not about the given details in a story but all the things that could be inferred from them. His work served to show specific knowledge required to answer such questions, giving a sense of the magnitude of the challenge. Work by Schank and his colleagues at Yale in the late 70s and early 80s (Cullingford, 1978; DeJong, 1982; Schank & Ableson, 1977; Wilensky, 1978) proposed specific classes of knowledge structures and showed how they applied to inferences about commonplace situations, causality, beliefs and intentions. This work was influential, but the use of rich knowledge in language understanding largely disappeared in the research community due to concerns about the lack of robustness and scalability (Lehnert, 1994) as well as the high knowledge engineering cost for applicable world knowledge.

The statistical revolution in natural language research shifted the focus almost entirely towards shallow, well defined phenomena such as parse trees, word sense disambiguation, entity recognition and semantic role labeling. This has been successful in creating clear evaluations for clear problems that focus the community and make for measurable progress. However, that progress has been detached from the larger goal of language understanding, realizing the danger of losing the forest for the trees. Not only are these tasks knowledge-poor, they are generally disconnected from any pragmatic context that might constrain or inform their processes. This lack of context makes knowledge that much harder to apply to problems and can exacerbate the perception that it cannot be done. In this environment, narrative understanding, with its need for rich knowledge and deep semantics, has seen little attention.

In other fields such as linguistics, discourse psychology and philosophy there are a number of broad theories of language use (Asher & Lascarides, 2003; Grice, 1975; Wilson & Sperber, 2004) and narrative understanding (Graesser, Singer, & Trabasso, 1994; Ryan, 1992; Trabasso, Secco, & van den Broek, 1984) that stress the importance of world knowledge and pragmatic context. However, they are either defined at such a high level that they provide little specific formalism, or they present a well-developed theoretical framework that has not been tested on narratives beyond hand-crafted examples. Instead, they follow the well-established methodology of demonstrating the capability to capture phenomena on example pairs of sentences considered in isolation. This is an effective way to identify and begin to explore these phenomena, but there have been few working efforts to move towards realizable computational models that address real narratives.

I believe that it is important for the natural language field in artificial intelligence to expand our view to encompass the entire task of language understanding. Human language competence is just as dependent on deep semantics, pragmatics and world knowledge as it is on lexical-syntactic concerns. But while the limits of shallow understanding are being thoroughly explored, the impact of those higher-level factors is not. Likewise, in the heyday of deep understanding research, the subtleties of lexical-syntactic presentation (e.g. word choice, parallel surface forms, repetition) were underappreciated. To bring them together I suggest that we place more value on research that addresses reasoning tasks that require world knowledge in a plausible, pragmatic context with real linguistic artifacts. There have been some such projects in the last few decades in the field of language understanding. Allen's work on dialogue (J. Allen et al., 2007; J. F. Allen et al., 2001), Hobbs' work on mechanical diagnosis (Hobbs, 1986) and Wilensky's work on automating online help (Wilensky et al., 2000) integrate world knowledge and pragmatic concerns into language processing. In doing so, they point the field towards using the wealth of shallow processing techniques for the actual goal of language understanding.

In this work, I apply this approach to the problem of narrative understanding. Rather than creating input examples tailored to particular phenomena, I address textual artifacts not created for the purpose of this research. I study written stories in their entirety to avoid arbitrary discourse segmentation and ill-defined discourse context common in linguistic work. To provide a well-defined context, I rely on task-oriented models of narrative pragmatics. Understanding a narrative in the most general sense is highly subjective and very difficult to model or evaluate. However, there is no shortage of plausible, knowledge-rich reasoning tasks, grounded in narrative texts, which can be used to provide pragmatic context. This work demonstrates that it is possible to bridge the gap from natural language to knowledge-rich reasoning tasks using widely available resources and established approaches in a novel combination. I show that it can accomplish novel, difficult inferential tasks over natural language narratives. Finally, I argue that the research on the reasoning side is enhanced by grounding in language at the same time that the research on the language side is enhanced by the context of the task.

1.1 Task-oriented models of narrative pragmatics

In this dissertation I consider three models of novel reasoning tasks for narrative understanding. Each reasoning task requires combining lexical-syntactic processing with deep semantics and world knowledge, and the model of each task defines a pragmatic account of understanding: those factors found or inferred in the narrative that have an impact on the outcome of the reasoning.

Two of these reasoning tasks are taken from research in *cognitive modeling*. The first involves making judgments about the proper course of action in morally-laden tradeoff scenarios, while the second involves assigning blame to agents in scenarios describing their involvement in a negative outcome. Cognitive modeling is based on the hypothesis that cognitive processes can be modeled as computation. Theses formal computational models of psychological theories rigorously explore the details and assumptions of those theories and can be evaluated against human performance. Many cognitive simulations are grounded in narrative texts, but due to the lack of language understanding systems up to the task, those texts must be manually encoded by experimenters in formal representations. This process is labor-intensive and error prone and leads to the problem of *tailorability*, since the simulation authors (or people working closely with them) do the encoding. Because they are based on psychological theories, these cognitive models give a very precise account of the factors in the text that impact reasoning. They also inherit the context of those theories, including careful circumscription of prior knowledge and pragmatic goals. This makes them a particularly attractive class of reasoning tasks to be integrated with language understanding.

The third reasoning task explores the notion of the meaning of a narrative. Taking the view that narrative is an intentional communication, the narrator is assumed to have some aim in telling the story which he or she believes is relevant to the audience. By social contract, the audience expects to find one or more such relevant meanings in the narration. These meanings may or may not be the same, and their similarity may be taken as one measure of the effectiveness of the communication. To summarize the vast amounts of study on this topic, I will quote from the notable narratologist Prince who states, "…understanding a narrative is not only being able to summarize it and paraphrase it in certain ways or to answer certain questions about its content; it is also (and perhaps even more so) being able to give an account of its "message", describe what (more or less) general subject or truth it illustrates, specify what "it is getting at", put forth its "point"." (Prince, 1983) Prince goes on to describe the need to better understand the pragmatic concerns that lead to inferring this point, acknowledging that narrative pragmatics is "…anything and everything that seems to be pertinent to narrative and that we do not have very clear ideas about" (Prince, 1983). Some clarity about the pragmatic concerns of narrative can be gained from Grice's maxims, which apply to all intentional communication (Grice, 1975), and from the work of Labov on how people tell personal stories. Labov suggests a function of *evaluation* where the narrator follows certain conventions to identify for the hearer the most significant elements of the narrative (Labov & Waletzky, 1966). In this work I explore the hypothesis that an inferential task where the reader expects to infer elements such as goals, threats and outcomes as well as contrasts, parallels and commentaries can serve as such an evaluative mechanism. By performing this task, the reader heuristically disambiguates elements of the narrative and brings attention to the likely meanings being conveyed. I test the effectiveness of this by showing that the system can match a set of fables to their morals.

These tasks require an approach that accounts for deep semantics, world knowledge and pragmatic goals to infer implicit content from the surface form of the input texts. To address this challenge I propose a practical approach that makes necessary concessions to the complexity of the problem without sacrificing those constraints.

1.2 Claims and contributions

The first contribution of this dissertation is an implemented approach to language processing that constructs deep, formal representations of narrative text which are suitable for pragmatically driven, knowledge rich reasoning tasks. Specifically, these representations are 1) in a standard logical form, 2) grounded in a large-scale ontology and 3) able to capture significant semantic distinctions common in narrative text. The interpretation process is a novel integration of existing techniques and resources that is guided by the pragmatic concerns of the reasoning task. Sentence-level processing uses *compositional frame semantics* to combine knowledge-rich *subcategorization frames* (Fillmore, 2006) with a limited grammar and an optional user intervention model. This efficiently generates highly expressive sentence-level representations that delay ambiguity resolution with explicit *choice sets*. These representations are automatically converted into *discourse representation structures* (Kamp & Reyle, 1993) to support dynamic discourse update and query-driven reasoning at the discourse level. This allows task pragmatics to guide contextual interpretation, including disambiguation, using *limited evidential abduction*. I show that this approach is empirically viable despite relying on higher-order predicate calculus and abductive reasoning to gain expressive power. I have evaluated the implemented system, *Explanation Agent (EA) NLU*, by showing that it is sufficient for three different reasoning tasks, over five evaluations spanning four different sets of stories. This supports the claim that this approach to narrative understanding is broadly applicable and not tied to a particularly narrow class of story or type of reasoning.

The second contribution of this dissertation is a theory of narrative pragmatics that defines a set of reader expectations that can be used as a heuristic guide for understanding. I cast these expectations as an inferential task that, unlike the more specific reasoning tasks I have studied, allows abductive disambiguation to be done incrementally (i.e. sentence by sentence). I claim that recognizing these expectations in this manner affords a sufficient understanding of some non-trivial classes of meanings communicated in narratives. I have implemented this theory as a pragmatic task for EA NLU and demonstrated a sufficient understanding of a set of Aesop's fables to identify the best-match moral for each fable.

1.3 Organization

Due to the breadth of topics covered in this dissertation, each chapter contains a background and a related work section.

Chapter 2 focuses on the practical EA NLU approach and implementation. Background and related work in platforms for knowledge-rich reasoning and inferential language understanding are given. This chapter provides evidence that the use of this system provides a consistent, principled translation from natural language to formal representation.

Chapter 3 describes the use of EA NLU with two cognitive models, one for moral decision making and one for the attribution of blame. Background and related work on the cognitive models and other approaches to knowledge-rich reasoning over text are given. Experiments with moral dilemma scenarios, corporate program scenarios and Iranian folktales are presented. This chapter provides evidence that EA NLU and the practical language understanding approach are sufficient to facilitate the use of natural language input for novel reasoning tasks.

Chapter 4 describes limited evidential abduction as a general-purpose reasoning mechanism for combining the linguistic disambiguation task with general task models. Background and related work on abductive reasoning and its prior application to language understanding are given. An experiment using abduction to automatically disambiguate the moral decision scenarios is presented. This chapter provides evidence that limited evidential abduction is sufficient for the disambiguation task and does not scale poorly with key complexity factors.

Chapter 5 describes a theory of narrative pragmatics that is used as a guide to incremental abduction in a much more general understanding task. Background and related work in narrative theory and computational story understanding are given. An experiment using this theory to identify appropriate morals for a set of Aesop's fables is presented. This chapter presents evidence that narrative pragmatics can capture a facet of understanding the meaning of a story, and that the EA NLU system is capable of supporting such an investigation.

Chapter 6 summarizes the claims of this dissertation and discusses some future directions.

2. Practical language understanding in EA NLU

This chapter describes my practically motivated approach to language understanding and its implementation in EA NLU. The constraints and goals of this approach are informed by the challenge of facilitating natural language input to knowledge-rich reasoning tasks. In particular, I focus on cognitive models which provide a clear task and a detailed account of understanding necessary for that task. Cognitive modeling experiments provide a novel venue for natural language work.

The foremost requirement for a semantic translation system seeking to support cognitive modeling experiments is sufficient *semantic breadth*. A high-level definition of semantic breadth is what the system can understand: the number of distinct scenarios, expressed in natural language, which can be translated by the system into distinct formal representations suitable for inference. However, this requires considerably more specification to be a useful implementation goal, and cognitive models provide that. For any given model there are sets of input stimuli, and within those sets there are salient similarities and differences that are hypothesized to result in predictable variations in reasoning and/or response. For example, in the moral decision making scenarios discussed in detail in the next chapter, each scenario contains a choice to intervene or not that is presented to the reader. One salient difference between the scenarios is the agent/patient roles of the proposed intervention action. A translation system is useful for this model only if its representations capture the presence of the choice and the different role assignments expressed in the different scenarios. Thus semantic breadth can be stated as a specific measure of the coverage of a translation system with respect to multiple cognitive models. To cover a model, the system must be able to translate the salient aspects of stimuli for that model. Further, the reasoning environment for the representations must be sufficient to implement that particular cognitive model.

The requirement of high semantic breadth raises a number of challenges for computational NLU. First, it requires rich, large-scale world knowledge that grounds types and relations in a consistent reasoning framework. The system must have some notion, for example, of how a fight is different than a hug and how causing is different than intending. This can be worked out in terms of particular models, which is a useful approach, but it must answer the question of how it will scale to unknown models. Second, it requires higher logical expressiveness than standard first-order logic (FOL) provides. Extensive work in the logic of language demonstrates that FOL is insufficient for capturing the range of ideas that can be communicated in natural language cf. (Boolos, 1984). Third, the expressiveness of the representation language is moot if the interpretation process cannot translate that range of ideas.

The state of the art of natural language understanding is such that there must be limitations in any implemented system. The challenges of facilitating natural language input to cognitive modeling experiments eliminate the very common strategy of limiting semantic breadth by working in a limited domain, working with a fixed corpus or reducing the expressiveness of the formal representations. This work supports the claim that a novel integration of knowledge-rich *subcategorization frames* (Fillmore, 2006), *compositional frame semantics*, explicit *choice sets*, *discourse representation structures (DRS)* (Kamp & Reyle, 1993) and query-driven back-chaining are together an effective approach to natural language understanding for cognitive modeling. I contribute a detailed, practical approach to knowledge-rich language understanding, and an implemented system, *Explanation Agent (EA) NLU*, both to test that claim and as a tool for cognitive modeling experiments.

In this chapter I will describe the implementation of this approach. I start with an overview of how this approach addresses the challenges I have described. I then discuss other research that this work builds on, followed by the details of the implementation. Specifically, how compositional frame semantics for sentence-level interpretation combined with query-driven proof of discourse-level facts implement this approach in the EA NLU interpretation process. Finally, I contrast related approaches and conclude with a general discussion. Evaluation against specific cognitive models is found in chapter 3.

2.1 Practical language understanding

The challenge of supporting knowledge-rich reasoning tasks across an unrestricted number of domains requires large-scale world knowledge. A great deal of research has investigated how far computational systems can get without such knowledge (due to the difficulty of constructing it). Here I take the view that such knowledge must be obtained, and that current, practical, large-scale knowledge resources can be effectively used. This implementation uses a knowledge base consisting of the contents of ResearchCyc plus our own extensions, together around 2 million facts at present. This provides not only world knowledge and a large-scale ontology, but also extensive, formal links between lexical terms and logical forms that are grounded in the ontology and axiomatized for reasoning. All this knowledge is expressed using the CycL language. CycL is a general-purpose higher order predicate calculus, allowing it to support a wide range of complex conceptual forms. In particular, it supports numerical and logical quantification, modal operators, higher-order relations and the use of microtheories as logical contexts. More details on ResearchCyc and CycL are given in section 2.2.1. These resources address the need for large-scale knowledge, connected to language in an expressive environment. However they raise three additional concerns: how to translate complex linguistic constructs using already complex semantic translations for terms, how to apply the world knowledge to the interpretation process and how to control the complexity of the process.

To aid in controlling complexity, I use a limited English grammar. In contrast with this approach, most recent explorations of automatic interpretation have focused on maximizing *syntactic breadth.* The domain of surface forms is greatly expanded, but at the cost of using impoverished internal forms. That is, a limited number of internal forms can be expressed in a wide variety of ways. Having multiple ways to say the same thing makes an NL system easier to use (i.e. increase habitability, cf. Haas & Hendrix, 1980), but the goal here is to maximize semantic breadth and the current state of the art does make this a trade off. In this approach, I use a grammar which supports at least one surface form for each internal form that it is capable of generating. Additional surface forms are often supported due to the compositional nature of the grammar, but only for user convenience. Again, it is not the number of surface forms that can be parsed that matters here, but the number of surface forms that can be parsed and interpreted as distinct semantic forms suitable for further reasoning.

The semantic interpretation process in this approach is divided into two levels. Compositional frame semantics are used at the sentence-level and query-driven back-chaining is used at the discourse-level. For each new sentence in a discourse, compositional semantics provide a fast, efficient way to build complex semantic representations from knowledge-rich subcategorization frames. Because compositional semantics factors out context in the composition of each syntactic constituent, it is able to handle nested constructs without becoming computationally intractable. Ambiguities are generated but maintained in packed representations for later disambiguation. The resulting forms from this composition are transformed into DRS which allow the logical form of each sentence to incrementally update a logical form for the discourse as a whole. This update process is based on query-driven back-chaining, where a particular reasoning task queries for facts to support its reasoning. This general-purpose reasoning allows pragmatic context and world knowledge to guide and constrain the discourse-level interpretation. However, deduction in higher-order reasoning is unsatisfiable, and the complexity of this process is dependent on the axioms of the reasoning task. By separating this process from the sentence-level composition, this approach mitigates that cost rather than placing constraints on the types of reasoning allowed.

2.2 Background

2.2.1 Cyc

The Cyc project (Lenat & Gupta, 1990) has worked for over twenty years building a knowledge base formalizing a broad selection of common-sense background knowledge. One of the primary goals of this project is to provide knowledge that is suitable for multiple reasoning and problem-solving tasks across many domains. It is based on the idea that systems limited to special-purpose domain knowledge, while effective in many tasks, are limited by brittleness and difficulty extending to unforeseen problems. In contrast, Cyc aspires to broad coverage that will ultimately support unforeseen future knowledge representation and reasoning tasks.

The ResearchCyc KB, available to the research community, currently contains more than 5 million assertions (facts and rules) (Curtis et al., 2009). This knowledge is represented in CycL, a higher-order logical language based on predicate calculus. CycL has a number of higher-order features, including quantification over predicates, functions, and sentences, and the ability for predicates to take predicates as values, including themselves in some cases (Ramachandran, Reagan, & Goolsbey, 2005). Higher-order predicates are used for logical relations (e.g. *and, or*, *implies*), quantification (e.g. *forAll*, *thereExists*) and modality (e.g. *willBe*, *possible*). Every assertion in CycL occurs in the context of a *microtheory*, allowing for the representation of competing theories and claims. Microtheories can generalize from one another providing for inheritance among contexts.

Much of the power of Cyc comes from a large-scale ontology. *Collections* within the ontology represent a kind or type of thing whose instances share a certain set of properties. The generalization relation *genls* is used to relate collections to one another while the membership relation *isa* relates them to instances (which may also be collections). Constraints (such as argument types for a predicate) and general axioms within the knowledge base are expressed in terms of this ontology to increase their effective power. Predicates themselves are organized in a similar fashion with the relation *genlPreds* in order to support inheritance of constraints and axiomatization. Creating knowledge within the framework of this ontology both increases consistency and multiplies effort.

2.2.1.1 Events and roles in Cyc

The knowledge in Cyc heavily uses Davidsonian representations of events (Davidson, 2001). Reified events belong to collections of such events, which generalize to the collection *Event* and through that to the collection *Situation*. Other features of the event are expressed using binary *role relation* predicates. For example, the representation of an event where a cake is eaten would contain the assertions:

(1) (isa eat1234 EatingEvent)

(2) (isa cake1234 Cake)

(3) (consumedObject eat1234 cake1234)

where *EatingEvent* and *Cake* are collections and (3) is a role relation between the eating event and the object being eaten.

2.2.1.2 Subcategorization frames and frame semantics in Cyc

ResearchCyc includes linguistic knowledge relating surface forms in natural language to formal logic forms using concepts in the ontology. The concept of a word belongs to the collection *LexicalWord*, and by convention is represented as its base form with "-TheWord" appended onto it. Thus the English word "car" is represented by the concept *Car-TheWord*. Lexical knowledge about this word is represented in terms of this concept. For example, valid parts of speech are expressed using the predicate *posForms* which relates a *LexicalWord* to a *SpeechPart* as in:

(4) (posForms Car-TheWord CountNoun)

Semantic knowledge in Cyc is based on extensive *denotations* and *subcategorization frames*. Denotations map directly from lexical terms to concepts within the Cyc ontology. They are stored as assertions of the form:

(5) (denotation <LexicalWord> <SpeechPart> <sense> <concept>)

where the sense number is unused and the concept is any concept in the knowledge base. In our example, the word "car" can be used to denote the concept of an automobile:

(6) (denotation Car-TheWord CountNoun 0 Automobile)

Subcategorization frames in Cyc follow Fillmore's theory of frame semantics (Fillmore, 1982) and map from a term to a logical form. Subcategorization of a word selects for the syntactic roles that it expects. A transitive verb, for instance, expects to apply to a subject and an object. Subcategorization frame types and the roles they expect are related in Cyc through the predicate *subcatFrameKeywords*. In the case of a transitive verb expecting a noun phrase complement, the assertion is:

(7) (subcatFrameKeywords TransitiveNPFrame :OBJECT)

Frame semantics suggest that for a given use of a term, a semantic frame can be constructed that expresses a logical representation of that usage in terms of syntactic roles. For example, in the sentence:

(8) I ate the sandwich.

the verb "ate" is being used transitively to indicate that an event took place, performed by a particular actor who consumed a particular object. There is an :ACTION role filled by the reified event, a :SUBJECT role filled by the actor "I" and an OBJECT role filled by the entity "the sandwich". The semantic frame for this usage is:

(9) (and (isa :ACTION EatingEvent)
(performedBy :ACTION :SUBJECT)
(consumedObject :ACTION :OBJECT)))

Subcategorization frames in Cyc are asserted using predicates of the type *SemTransPred.* These specialize on a *SpeechPart* and express the relation between a *LexicalWord*, a subcategorization frame type and a semantic frame. The assertions are of the form:

(10) (<semTransPred> <LexicalWord> <sense> <FrameType> <semantic frame>)

where sense is ignored and the particular *SemTransPred* is connected to a *SpeechPart* by an assertion of the form:

(11) (semTransPredForPOS <SpeechPart> <SemTransPred>)

Thus the relevant assertions for the example subcategorization frame for sentence (8) are:

(12) (semTransPredForPOS Verb verbSemTrans)

(13) (verbSemTrans Eat-TheWord 0 TransitiveNPFrame
(and (isa :ACTION EatingEvent)
(performedBy :ACTION :SUBJECT)
(consumedObject :ACTION :OBJECT)))

2.2.2 COMLEX

The COMLEX syntax project (Grishman, Macleod, & Wolff, 1993) created a “moderately broad coverage” lexicon to be shared with the automatic language research community through the Linguistic Data Consortium. COMLEX is a straightforward dictionary with entries for English words, phrases and punctuation. Each entry connects a lexical term to a set of lexical-syntactic features, each having a label and one or more values. The *root* feature in every entry maps the term to a base form, and multiple forms of the same word (e.g. tenses, plurals, superlatives) each have a unique entry. Lexical category (part-of-speech), agreement and verb form are other important features in the lexicon. The focus of the COMLEX lexicon is on providing detailed syntactic specifications in the form of feature classes that can be used to constrain parsing through subcategorization. For example, an adverb may have the *modif* feature *pre-adj* indicating that it may be placed prior to an adjective that it modifies (e.g. "*very* angry"). Similarly, verbs are subcategorized as to whether they can accept a direct object noun phrase, prepositional phrase, infinitive complement and so forth. Because these constraints are explicitly assigned to each form of each term, COMLEX provides a very powerful, highly detailed level of syntactic control. EA NLU uses COMLEX version 3.1, consisting of roughly 86000 dictionary entries. For additional details, see the reference manual (Macleod, Grishman, & Meyers, 1998).

2.2.3 EA NLU

The EA NLU system was originally developed by (Kuehne, 2004) to explore understanding descriptions of physical phenomena in natural language text. He demonstrated that common descriptions of such phenomena effectively mapped to a frame-based representation of *Qualitative Process Theory* (Forbus, 1984). Kuehne integrated Allen's bottom-up chart parser (J. F. Allen, 1994) with the COMLEX lexicon (Grishman et al., 1993) and subcategorization frames from Cyc. He also developed the *Qualitative Reasoning Group Controlled English (QRG-CE)* grammar used by the parser. The use of a limited-syntax language to factor out parsing difficulties was inspired by both CMU's KANT project (Nyberg et al., 2002) and Boeing's controlled language work (P. Clark, Harrison, Jenkins, Thompson, & Wojcik, 2005). Like these simplified languages, QRG-CE restricts grammar but does not *a priori* restrict the vocabulary. This enables most extensions to be made by adding vocabulary rather than changing the grammar. That version of QRG-CE had limited support for general syntax. Because the focus of the work was on identifying specific syntactic patterns, the grammar needed only to support those enumerated patterns and the sub-constituents that comprised them. Further, since those sub-constituents only needed to compose into the known patterns, they could largely be non-recursive, greatly reducing the possible combinations and the associated complexity. A later effort extended the grammar to support a set of question forms using the same control strategy, with the same limitations.

2.3 Compositional frame semantics

This practical language understanding approach uses two layers of complementary interpretation processing. At the sentence-level, compositional frame semantics are used in order to make computational complexity more predictable and less explosive. Arbitrarily complex logical forms are allowed in the semantics, but based on an actual set of usable semantic frames rather than a theoretical set. Additional control comes from limiting the syntactic forms via the simplified grammar. The guiding principle for the sentence-level processing is that ambiguous semantic distinctions must be preserved but do not need to be resolved. This enables the use of a compositional approach where the semantics of any phrase is dependent only on the semantics of its constituent parts. This limits sources of combinatorial explosion to a set of well-defined interactions that I return to at the end of this section.

2.3.1 Parsing architecture

Like the previous version, EA NLU uses Allen's bottom-up chart parser (J. F. Allen, 1994) with the COMLEX lexicon (Grishman et al., 1993) and the QRG-CE simplified English grammar. Text input is processed one sentence at a time and converted to a sequence of symbols, including

COMLEX symbols for punctuation (e.g. punc-comma, punc-question-mark). Each symbol is looked up in the lexicon, which returns either a null symbol or one or more nested list structures containing a set of lexical features. Those features represent a particular lexical-syntactic sense of the term. The result of the lookup is used to create one or more leaf constituents based on lexical category. In the case of a null result, an *unknown*-type constituent is created. Each leaf constituent is assigned a randomly generated *discourse variable* which is the reification of the conceptual contribution of that constituent. The leaf constituents are added to the chart and arcs are created or extended based on the grammar rules. The result of this process is one or more parse trees consisting of a root constituent and its sub-constituents.

2.3.2 Qualitative Reasoning Group Controlled English (QRG-CE)

The QRG-CE grammar is a limited subset of American English based on the earlier work of (Kuehne, 2004). In this work I have recreated the general syntax from scratch to make it considerably broader in application to narrative text. The grammar composes noun phrases, verb phrases, prepositional phrases (PP), adjective phrases (ADJP) and adverb phrases (ADVP) in a manner similar to Allen's textbook grammar (J. F. Allen, 1994) or other similar efforts (Jurafsky & Martin, 2009). Support for punctuation, sentence-level phrases (SLP) (including inversions for questions and passive voice), and coordinating conjunctions are likewise straightforward. I will not go into great detail on these parts of the grammar. It is worth noting that the verb phrase support includes auxiliary support for all tenses and aspects, and that coordinating conjunctions are supported between NP, VP, ADJP, ADVP and SLP constituents. The high-level patterns for physical processes and question answering from Kuehne's work were re-added to this new grammar. The entire QRG-CE grammar is available in Appendix A.

The syntactic extensions to QRG-CE made in this work are fairly standard and well understood, as are the basic semantic compositions. The contribution of this work is in using these techniques with highly expressive, knowledge-rich semantic frames to support higher-order compositions commonly found in narrative text.

2.3.3 Retrieving semantic knowledge

The semantic interpretation of a constituent in the parser is contained in the *sem* feature slot of that constituent. This slot is filled for new leaf constituents by retrieving semantic knowledge, in the form of denotations and subcategorization frames, as explained in section 2.2.1.2, from the knowledge base. The retrieval is based on the root form of the lexical term combined with the lexical category, both provided by COMLEX. The root form is converted to a Cyc *LexicalWord* concept by creating a symbol made of the root form with "-TheWord" appended to it. The COMLEX lexical category is mapped to a set of Cyc *SpeechPart* concepts by retrieving facts from the knowledge base of the form:

(14) (synonymousExternalConcept <SpeechPart> COMLEX31Lexicon <lexcat>)

For the *LexicalWord* and each *SpeechPart* derived for a leaf constituent, a set of zero or more denoted concepts and semantic frames are retrieved. The denoted concepts are used to generate additional semantic frames. In the general case, the denoting term is being used to express that an entity belongs to a collection. In this case the denotation of, for example, the concept *Bank* results in the semantic frame:

(15) (isa :NOUN Bank)

Other cases arise because denotations in ResearchCyc are not limited to collections. There is considerably more ambiguity in the denotation of a relation or function. Following from the principle of preserving ambiguity, these denotations result in the ambiguous semantic frames:

(16) (denotesRelation-Underspecified :NOUN <Relation>)

(17) (denotesFunction-Underspecified :NOUN <Function-Denotational>)

Leaf constituents may also be generated by recognition of proper names and common phrases. Proper names in EA NLU are implemented as a simple dictionary mapping words or phrases to Cyc concepts. These are retrieved in much the same way as denotation facts, but the resulting constituent uses the retrieved concept as both the semantics and the discourse variable. Thus the phrase "John Adams" results in a constituent with the discourse variable *JohnAdams*, an individual concept representing the second president of the United States.

Common phrases are stored in several ways in the knowledge base. ResearchCyc includes four predicates for specifying multi-word phrase semantics. Examples of the four types of expressions are:

(18) (multiWordString (TheList "boiling") Point-TheWord CountNoun BoilingPoint)

(19) (multiWordSemTrans (TheList "boiling") Point-TheWord CountNoun
GenitiveFrame (boilingPoint :POSSESSOR :NOUN))

(20) (compoundString Pay-TheWord (TheList "cash") Verb PayingWithCurrency)

(21) (compoundSemTrans Pay-TheWord (TheList "cash") Verb
IntransitiveFrame (and (isa :ACTION PayingWithCurrency)
(payer :ACTION :SUBJECT)))

The *multiWord* predicates correspond to cases where the multi-word phrase should inherit the lexical features of the last word in the phrase, in the examples the word “point”. By contrast, the *compound* predicates are used for phrases that inherit from the first word, in the examples the word “pay”. This allows the system to utilize this knowledge for different forms of the phrases, for example “boiling point” vs. “boiling points” and “pay cash” vs. “paid cash”. For each of these two types of predicates, there is the *-String* version and the *-SemTrans* version. These correspond to denoting a single concept, *BoilingPoint* in the first example, and providing a semantic translation frame, *(boilingPoint :POSSESSOR :NOUN)* in the second example.

The semantic frames retrieved from all these sources combine to form the set of possible semantics for a leaf constituent.

2.3.4 Lambda-composition and quantification

The Allen parser supports lambda-calculus composition of semantic features through role binding and replacement. Grammar rules specify feature constraints on sub-constituents as well as the resulting features on the composed constituent. EA NLU utilizes lexical-syntactic features to constrain the parsing process and enforce the simplified grammar. The grammar rules take the form:

(22) (<constituent> <name> <sub-constituent 1> … <sub-constituent n>)

For example, a basic rule for creating a sentence-level phrase (SLP) from a noun phrase (NP) and a verb phrase (VP) has the form:

(23) ((slp (var ?varvp) (agr ?a) (:SUBJECT ?varnp) (sem (and ?semnp ?semvp)))
-slp->np-vp-
(np (var ?varnp) (agr ?a) (sem ?semnp))
(head (vp (var ?varvp) (agr ?a) (sem ?semvp))))

The name of the rule is *–slp->np-vp-*. The SLP constituent inherits the variable of the *head* sub-constituent, in this case the VP. The agreement feature of the VP and NP sub-constituents must match and is inherited by the SLP. The semantic feature of the composed constituent is specified as the conjunction of the semantic features of the sub-constituents. Simple conjunction is the most basic form of semantic composition used in QRG-CE. In this example, the variable of the NP sub-constituent has been bound to the semantic role feature :SUBJECT. As part of the semantic composition process, role keywords are replaced with any such bound role features. Thus, if the semantic feature for the VP sub-constituent, with discourse variable *eat1234*, is:

(24) (and (isa eat1234 EatingEvent) (performedBy eat1234 :SUBJECT))

and the semantic feature for the NP sub-constituent, with discourse variable *him1234*, is:

(25) (isa him1234 Person)

then the semantics of the composed constituent, with discourse variable *eat1234* inherited from the VP sub-constituent and :SUBJECT feature *him1234*, will be:

(26) (and (isa eat1234 EatingEvent) (performedBy eat1234 him1324) (isa him1234 Person))

Frame semantics provide a broad foundation of rich semantic forms that can be constructed from language. Events and role relations alone represent a considerable semantic breadth. However, the real breadth of natural language begins to come out in the composition of such events, relations and other assertions into complex forms. As a controlled language, QRG-CE can be viewed as a specification of a set of such compositions. These compositions must be independent of the content of the frames being composed: a frame describing a particular event is valid regardless of where in the composition it falls. This requirement maximizes the semantic breadth the system can achieve.

2.3.5 Logical and numerical quantification

One of the significant challenges for semantic interpretation is proper quantification of the entities being described. Frame semantics supports quantification is two ways. First, the semantic role slots in the frame specify knowledge about expected entities. This assumes that the entities being referenced will be created and quantified by the compositional process, independent of the particular frames being invoked. Second, because the frames allow any valid logical expression, explicit quantification can be included in the translation. Frame-independent existential quantification takes place at the composition of NP and VP constituents in QRG-CE. The discourse variable for the noun or verb head in the phrase is existentially quantified over the composed semantics. The specification for the semantic feature of the composed constituent looks like (for a simple NP):

(27) (thereExists ?varnoun ?semnoun)

Thus, a NP such as “the bank” takes the noun constituent semantic translation:

(28) (isa :NOUN Bank)

and composes the NP constituent semantic translation:

(29) (thereExists bank1234 (isa bank1234 Bank))

When multiple, quantified sub-constituents are composed, it is necessary to resolve scoping of the quantifiers. In the simplest case of a conjunction of multiple existentially quantified expressions, the quantified variables taken together are scoped over the conjunction of quantified facts. To repeat the above example in expressions n-n with quantification, if the semantic feature for the VP sub-constituent, with discourse variable *eat1234,* is:

(30) (thereExists eat1234
(and (isa eat1234 EatingEvent) (performedBy eat1234 :SUBJECT)))

and the semantic feature for the NP sub-constituent, with discourse variable *him1234*, is:

(31) (thereExists him1234 (isa him1234 Person))

then the semantics of the composed constituent, with discourse variable *eat1234* inherited from the VP sub-constituent and :SUBJECT feature *him1234*, will be:

(32) (thereExists (TheList eat1234 him1234)
(and (isa eat1234 EatingEvent) (performedBy eat1234 him1324)
(isa him1234 Person))

Frame-independent universal quantification is supported by QRG-CE through the use of the specific quantifiers “all” and “every” applied to common noun phrases (CNP) to create an NP. In the same way an existential expression is constructed, the discourse variable for the noun head in the phrase is quantified over the composed semantics. The specification for the semantic feature looks like:

(33) (forAll ?varnoun (implies ?semnoun :SCOPED-CLAUSE))

Thus the NP "every dog" has the semantic translation:

(34) (forAll dog1234 (implies (isa dog1234 Dog) :SCOPED-CLAUSE))

The binding for the semantic role :SCOPED-CLAUSE is determined by the use of the NP as a subject, object, oblique-object, etc.

Numerical and qualitative quantification is supported with reified groups and universal implications based on group membership. A NP such as "5 cats" translates to the logical form:

(35) (thereExists group-of-cat1234
(and (isa group-of-cat1234 Set-Mathematical)
(cadinality group-of-cat1234 5)
(forAll cat1234 (implies (member cat1234 group-of-cat1234)
(isa cat1234 Cat)))))

The discourse variable group-of-cat1234 is used as the variable for the constituent of this NP, allowing it to compose with other semantic features in the standard ways discussed above. This reification of the group also simplifies subsequent references. Subsequent references applicable to the generic individual cat1234 can be inserted into the antecedent of the implication using dynamic logic principles discussed below. Qualitative quantification, such as *many cats* is handled in the same way, but the cardinality statement is replaced with a qualitative statement such as:

(36) (qualitativeExtent group-of-cat1234 Many)

Because the compositional process relies only on the semantics of the immediate sub-constituents, it does not matter whether quantification (or any other logical form) came from a semantic frame or was created by a prior compositional step. Thus quantification in a semantic frame is seamlessly composed with frame-independent quantification. For example, the term "friend" invokes a semantic frame which explicitly quantifies some agent:

(37) (thereExists ?AGENT (friends ?AGENT :NOUN)))

The frame variable ?AGENT is converted to a discourse variable (e.g. agent1234) and the quantification is composed with other expressions as described above.

2.3.6 Nested clauses

Existential quantification and conjunction are the most common compositional operations in QRG-CE. This is sufficient for expressing sets of states and events that occur in the same reality, such as a sequence of events in a simple story. To extend beyond this to hypothetical, conditional, negative and other modal expressions, QRG-CE uses nested clauses. That is, compositions where the semantic interpretation of a clausal constituent is treated as the argument to a higher-order predicate in its parent constituent. Over the four sets of stories studied in this work, I found that nested clauses occur multiple times in every story while anaphoric references over complex quantifier ambiguities, the subject of much work in linguistics (e.g. the "donkey anaphora sentences"), rarely occur.

2.3.7 Modal operators

QRG-CE supports modal statements using the Cyc modal operators *willBe* and *possible*. willBe is introduced by future tense verbs and operates over the semantics of the VP. In the moral decision scenario example, the phrase "will be extinct" has the semantic translation:

(38) (willBe (thereExists be1234
(and (isa be1234 Extinction) (objectActedOn be1234 :SUBJECT)))

In terms of semantic breadth, this creates a clear distinction that the existence of the Extinction event be1234 is a default future, not something that has occurred. The statement “Dinosaurs are extinct”, by contrast, would have no such nested structure.

Additional modal operators are introduced into an interpretation by semantic frames. COMLEX recognizes a set of auxiliary verbs as having the lexical feature *modal*. Those auxiliaries invoke modal semantic frames, such as for the term “can”:

```
(39)  (modalVerbSemTrans Can-TheWord 0 TransitiveBareInfinitiveFrame
            (possible :CLAUSE))
```

The :CLAUSE role is bound in QRG-CE not to discourse variables, but to the entire semantic translation of a sub-constituent. One grammar rule for applying the above frame (some lexical constraints removed for readability) is:

```
(40)  ((vp (agr ?a) (vform ?vform) (var ?varv)
          (:CLAUSE ?semv)
          (sem ?semaux))
      -vp->modal-vp-
      (aux (modal +) (sem ?semaux) (agr ?a) (vform (? vform pres past)))
      (head (vp (vform pres) (var ?varv) (sem ?semv))))
```

Note that the semantic translation of the VP sub-constituent is bound to the :CLAUSE role, and the semantic feature of the composed constituent consists of only the semantic translation of the modal auxiliary. The resulting form for “can save” as in the moral decision scenario example is:

```
(41)  (possible (thereExists save1234
                    (and (isa save1234 RescuingSomeone)
                         (performedBy save1234 :SUBJECT)
                         (beneficiary save1234 :OBJECT))))
```

Again, the existence of the RescuingSomeone event save1234 is not entailed by the model but rather the possibility of its existence. The ability to reason about hypothetical and default events is crucial in many narratives. An inevitable series of events is a very simplistic type of story. More typically, it is the actions taken by agents to cause or avoid particular futures, as well as their awareness and motivations, that make stories interesting. The moral decision scenarios and cultural folktales are built around events that happen with and without intervention. The corporate program scenarios involve discussion of future events, including discussion of causal outcomes, for the purpose of deciding what actions to take.

2.3.8 Negation

A similar challenge is raised and addressed in the case of negated events and propositions. In the case of a negated proposition, the obvious approach is to wrap the proposition in a modal *not* operator. This results in the phrase "not green" translating to:

(42) (not (mainColorOfObject :NOUN GreenColor))

However, it is a more complex problem to negation a quantified verb in Davidsonian representation. It is not sufficient to negate the *isa* statement as that would imply the existence of an event which is not of that type. Nor is it satisfying to introduce an additional predicate, something like *notOccurs,* to distinguish reified events that did not occur. This again implies the existence of a specific event that did not occur. This again argues for the necessary complication of supporting higher-order logic through CycL in our practical language understanding approach. It is far more consistent to negate the quantification of the verb such that the phrase "did not move" translates to:

(43) (not (thereExists move1234
(and (isa move1234 MovementEvent) (primaryObjectMoving :SUBJECT))))

2.3.9 Clausal verb complements

Several classes of terms in English translate to a higher-order logical expression. These terms invoke semantic frames that use the :CLAUSE role as an argument to a higher-order predicate. QRG-CE supports composition of these clausal frames based on lexical features provided by COMLEX, specifically the verb features *parenthetical* and *vsay*, verb subcategorizations allowing an infinitive complement and adjective subcategorizations allowing a that-clause complement.

Parenthetical verbs such as *believe* and *realize* expect a clausal complement, as in this semantic frame for *realize*:

(44) (and (realizedProp :ACTION :CLAUSE)
(isa :ACTION RealizingThat)
(doneBy :ACTION :SUBJECT)))

In the same manner as modal operators are supported, QRG-CE binds the entire semantic feature of the complement constituent to the :CLAUSE role. This results in a higher-order expression with the predicate *realizedProp* indicating what has been realized. Because arbitrary nesting is allowed in the composition, it is possible to express what is realized as a simple fact, a quantified event or entity, or even a modal statement about what will be, might be or is not. For example, the phrase “realizes he can not eat” generates a *realizedProp* of the form:

(45) (realizedProp realize1234
(thereExists he1234
(not (possible (thereExists eat1234
(and (isa eat1234 EatingEvent) …))))))

Verbs with the feature *vsay* expect a clausal complement indicating what was communicated. This may be implicitly stated, as in:

(46) He said they were there.

or explicitly stated as in:

(47) He said, "They were there."

QRG-CE supports both cases, resulting in a speaking event of some type (e.g. Speaking, Yelling) and a *infoTransferred* role connecting the event to the logical expression of what was communicated. The resulting translation is of the form:

(48) (thereExists (TheList he1234 say1234)
(and (isa say1234 Informing)
(senderOfInfo say1234 he1234)
(infoTransferred say1234
(thereExists (TheList they1234 there1234)
(objectFoundAtLocation they1234 there1234)))

A third class of verbs expects an infinitive complement, expressing a relationship between the verb and another event or action. Examples are "wanted to run" and "tries to get them". These verbs invoke *InfinitivePhraseFrames* that have the syntactic role :INF-COMP. For example, the verb *wanted* has the semantic frame:

(49) (desires :SUBJECT :INF-COMP)

A simplified rule for composing an infinitive complement phrase looks like:

(50) ((vp (var ?varv) (vform ?vform) (agr ?a)
(:INF-COMP ?semnp)
(sem ?semv))
-vp->vp-npinf-
(head (vp (vform ?vform) (subcat (? sc to-inf-sc to-inf-rs)) (agr ?a)
(var ?varv) (sem ?semv)))
(np (var ?varnp) (sem ?semnp)))

The *subcat* feature of the VP sub-constituent specifies that only VPs with a *subcat* feature containing either *to-inf-sc* or *to-inf-rs* are valid for this composition. The resulting translation is the translation of the VP with the translation of the infinitive NP sub-constituent bound to the :INF-COMP role. The resulting composition for *wanted to move* is:

(51) (desires :SUBJECT
(thereExists move1234
(and
(isa move1234 Movement-TranslationEvent)
(primaryObjectMoving move1234 :SUBJECT)))))

A class of adjectives also expects a clausal complement, specifically a *that-clause*. QRG-CE supports this syntax for adjective phrases that have a *subcat* feature containing either of the COMLEX concepts *that-s-adj* or *extrap-adj-that-s*. Examples are “sure that it smells” and “true that he took them”. The complement in this case is a SLP beginning with the adjunct adverb *that*. As in the above cases, the semantic translation of the SLP is bound to the :CLAUSE role which is expected to be used in the semantic frame for the adjective being complemented.

A second, more general use of the infinitive form arises when one action is done for the purpose of bringing about another. In the fifth sentence of the folktale example there is the phrase “run

… to inform the manager". The verb *run* does not subcategorize for an infinitive complement in COMLEX, and clearly that is not the intent here. Rather, the latter, infinitive verb expresses the purpose of the former verb. This is a common expression in the narratives I have worked with and thus an important breadth element. The relation between the two verbs is expressed as:

(52) *(purposeInEvent run1234 agent1234 (thereExists inform1234 ...))*

As with the other examples of clausal substitution, it is significant that the existence of the informing (which may be further qualified by modals or other operators) is wrapped up in the purpose. That is to say, the agent is running so that an informing event might exist.

In the case of this syntactic pattern where the former verb does subcategorize for an infinitive complement, the syntax is legitimately ambiguous: it might be expressing, for example, that an agent "wishes to eat" or that the agent "wishes *in order to* eat" (for some context where that makes sense). QRG-CE supports this type of legitimate ambiguity by allowing for both forms. This creates additional complexity in the parse trees, but this cost is mitigated by the compositional semantics. Further, because the ambiguity is preserved, contextual knowledge can be brought to bear on the disambiguation problem during the discourse-level processing. Sentence (53), taken from an Iranian folktale, is a good example of the type of semantic expressiveness that QRG-CE supports.

(53) *The first option is, he can try to run to the station to inform the manager*.

Note that the preposition "to" is used three times, to express three different relations with three different semantic transformations. The first, "try to run", is an infinitive complement which fulfills the clausal expectations of an *Attempting* action. The second, "run to the station", is a prepositional attachment indicating the location to which the translation indicated to the motion verb “run” is heading. And the third is the statement of purpose described above. This sentence is typical of the content of narrative scenarios. It is not just describing an event, but rather the possibility of an event for the purpose of another for the purpose of another detailed event. QRG-CE is designed to support this kind of semantic expressiveness by composing rich building blocks in a computationally tractable way.

2.3.10 Causal statements

The more general case of expressing causality is quite central to narrative communication. It is often tied up in the most significant point of a story to assert a certain causal understanding. QRG-CE supports this expression in two ways. The first is through use of the verb “cause” with an infinitive complement, as in:

(54) The opening would cause 2 species of fish to be extinct.

Importantly, the NP “2 species of fish” acts as the subject of the infinitive phrase “to be extinct”. This creates a legitimate ambiguity with transitive use of a verb with an infinitive complement as in:

(55) The ant used the branch to reach the shore.

In the latter case, the same syntactic structure is invoked, but the NP “the branch” acts instead as a direct object to the verb “used”.

The second syntactic pattern involves the subordinating conjunction “because”, as in:

(56) Because of a dam on a river, 20 species of fish will be extinct.

Trivially, the order of clauses may be reversed. Cyc defines a hierarchy of causal relations specialized for both arguments on the types *Thing*, *Situation* and *CycL-SentenceAssertible* (proposition). From a semantic interpretation standpoint, the distinction between thing and situation is not appropriate at this level. That is, the same syntactic pattern is used for a thing or a situation, and the resulting discourse entity is typed by *isa* assertions. Whether dam1234 is a thing or a situation is already known, making specialization of the causal relation an unnecessary increase in complexity. In contrast, the distinction between a reified entity and a proposition alters the structure of the semantic translation. Supporting causal relations between propositions is a significant increase in semantic breadth. Without such support, cause can only be asserted between discourse entities that occur in the model at the same level of nesting. That is, events that have actually occurred. A simple statement such as:

(57) The dog caused the accident.

can result in the quantified conjunction of three first-order assertions of the form:

(58) (thereExists (TheList dog1234 accident1234)
(and (isa dog1234 Dog)
(isa accident1234 Accident)
(causes-ThingSit dog1234 accident1234))

However, if one of the participants is hypothetical, as in the statement:

(59) Because of the dog, there will be an accident.

then the simple formalism is inadequate. Instead, EA NLU uses a nested higher-order statement to capture the relationship between propositions. In this example, the predicate *causes-*

ThingProp is used and the second argument receives a clausal binding (via the :CLAUSE role) such that the translation is:

(60) (thereExists dog1234
(and (isa dog1234 Dog
(causes-ThingProp dog1234
(willBe (thereExists accident1234 (isa accident1234 Accident))))))

Importantly, the resulting event *accident1234* is not incorrectly asserted to occur in the root context of the model.

2.3.11 Utterances

A particularly interesting form of nested clause is spoken utterance, whether expressed implicitly as in sentence (61) or explicitly as in sentence (62).

(61) She said the store is closed.

(62) She said, "The store is closed."

In both cases, the syntactic feature *vsay* in COMLEX is identified on the verb, allowing it to be followed by a SLP constituent (either quoted or not), in this case "the store is closed". In the former case, composition proceeds as in the other nested clause cases with the :CLAUSE role binding to the SLP. The applicable semantic frame for the verb "say" is:

(63) (and (isa :ACTION Informing)
(senderOfInfo :ACTION :SUBJECT)
(infoTransferred :ACTION :CLAUSE))

The resulting form asserts the things said in a nested context, within the *infoTransferred* expression, thus making the distinction between things globally asserted in the world and the assertion of things being spoken of in the world. The same semantic frame applies to the explicit utterance case, but the composition is complicated by the possibility of multiple sentences within the utterance as in:

(64) She said, "The store is closed. I will go tomorrow."

The bottom-up chart parser can certainly handle sentences as child constituents of other sentences. But that approach scales poorly because the numbers of possible parses for each quoted sub-sentence are multiplied together and there can be arbitrarily many of them. To avoid this problem, each quoted sub-sentence is independently parsed and the top level sentence contains only references to the semantic translation of them. This reference is represented in terms of sentence identifiers as in the expression:

(65) (infoTransferred say1234 (TheList Sentence-1234-1 Sentence-1234-2))

Representing the possible worlds of modalities, hypotheticals, negations and utterances as nested expressions has several advantages. The range of higher-order nesting predicates provides a great deal of semantic breadth, allowing for distinct reasoning about things that are, might be, were spoken of and so forth. At the same time, the content of a nested expression is the same regardless of the particular nesting or whether it is nested at all. Thus axioms that apply to the description of a man walking down the street can apply regardless of whether it is being observed, discussed, predicted, and so on.

2.3.12 Quantifier Scope

One of the complexities that arises from supporting general quantification and nested statements is ambiguity in quantifier scoping. EA NLU handles this in two ways. First, scoping ambiguities between existential and modal statements differ in the assumptions of existence. A statement such as:

(66) I will buy a house.

creates scoping ambiguity when the auxiliary “will” is composed with the VP “buy a house”. The resulting form is one of:

```
(67) (willBe
       (thereExists (TheList buy1234 house1234)
         (and (isa house1234 House-Modern)
              (isa buy1234 Buying)
              (buyer buy1234 :SUBJECT)
              (objectPaidFor buy1234 house1234))))

(68) (thereExists buy1234
       (willBe
         (thereExists house1234
           (and (isa house1234 House-Modern)
                (isa buy1234 Buying)
                (buyer buy1234 :SUBJECT)
                (objectPaidFor buy1234 house1234))))

(69) (thereExists house1234
       (willBe
         (thereExists buy1234
           (and (isa house1234 House-Modern)
                (isa buy1234 Buying)
                (buyer buy1234 :SUBJECT)
                (objectPaidFor buy1234 house1234))))
```

(70) (thereExists (TheList buy1234 house1234)
(willBe (and (isa house1234 House-Modern)
(isa buy1234 Buying)
(buyer buy1234 :SUBJECT)
(objectPaidFor buy1234 house1234))))

At issue here is whether the buying event and/or the house being bought pre-exist the default future state expressed by the modal *willBe.* Attempting context-independent disambiguation would not be consistent with the principles of this approach, so this ambiguity should be maintained. However, it is also the case that form (67) is true in exactly the cases where any of the other three forms are true. That is to say, it is true for all four forms that the buying event and the house exist at the time of the buying. Form (67) represents exactly and only the knowledge that can be assumed in this environment. Further assumptions, such as the pre-existence of the house, are not constrained by the surface form and can be made at a later time. Thus EA NLU resolves this ambiguity immediately and the composition results in form (67).

The second case of quantifier scope ambiguity is between existential and universally quantified statements. This is a well understood problem in computational linguistics and it suffices here to say that it represents legitimate ambiguity. This type of ambiguity is maintained as a choice set, described below, to be resolved based on context at the discourse-level processing.

2.3.13 Dynamic update with discourse representation structures

Natural language communication is inherently serial. As utterances are perceived and processed, internal knowledge must be updated incrementally. Standard logical forms are not well-suited to this type of update. Davidsonian event representations address this issue by allowing the flexibility to incrementally extend knowledge about an event. Logical assertions about an event can be made as information comes in, without the need to retract and change prior assertions. More significantly, higher-order statements can be made on the reified event itself or on some conjunction of facts about the event. For example, in the phrase “because he moved it”, the causal antecedent is not the event alone but the actor relation between the event and “he”. This is captured by the role relation:

(71) (performedBy move1234 he1234)

which can be used as an argument for a higher-order causal relation such as:

(72) (causes-PropProp (performedBy move1234 he1234) <consequent>)

Further, it may later be asserted that the moving had some other property – perhaps it was quick or careless – without altering the truth condition of the statement (72). By contrast, a causal statement made on the reified event:

(73) (causes-SitProp move1234 <consequent>)

has a different truth-condition that the event as a whole is the antecedent. This underspecifies the contribution of specific role relations to the cause, a condition that is commonly useful in natural language. The ability to make this distinction is a significant increase in expressiveness and thus in semantic breadth.

Discourse Representation Theory (DRT) (Kamp & Reyle, 1993) provides an analogous representation method for nested logical forms. Each level of nesting can be reified as a discourse representation structure (DRS) that represents an existential quantification over a conjunction of facts. These DRS can be embedded similarly to reified events and other entities. For example, the sentence:

(74) I did not swim.

generates the (simplified) compositional semantics in standard CycL form:

(75) (thereExists i1234
(not (thereExists swim1234
(and (isa swim1234 Swimming-Generic)
(providerOfMotiveForce swim1234 i1234))))

This translates to the nested DRS representation shown in Figure 1. Each existential quantification has been transformed into a DRS where the *universe* of the DRS is the set of quantified variables and the *conditions* are the set of facts quantified over. The nested DRS is denoted by the functional term *(DrsCaseFn DRS-1234)*, and that term is nested in the negation operator *not*.

Universe: i1234

(not (DrsCaseFn DRS-1234))

DRS-1234:

Universe: swim1234

(isa swim1234 Swimming-Generic)
(providerOfMotiveForce swim1234 i1234)

Figure 1. DRS for "I did not swim."

In EA NLU, the function *DrsCaseFn* is used to denote a DRS as a *case* (which is a subclass of a Cyc microtheory) with an identifying label. These DRS are implemented as cases using the binary relation *ist-Information* to express that a fact is true in a particular case. The universe of a DRS is expressed using the relation *variableInUniverse* to relate a discourse variable to a DRS identifier. The example in Figure 1, assuming that the outer DRS identifier is *DRS-root*, would be comprised of the facts:

(76) (variableInUniverse i1234 DRS-root)

(77) (ist-Information (DrsCaseFn DRS-root) (not (DrsCaseFn DRS-1234)))

(78) (variableInUniverse swim1234 DRS-1234)

(79) (ist-Information (DrsCaseFn DRS-1234) (isa swim1234 Swimming-Generic))

(80) (ist-Information (DrsCaseFn DRS-1234) (providerOfMotiveForce swim1234 i1234))

It has been shown that DRS is logically equivalent to predicate calculus and translations between the two are straightforward (Kamp & Reyle, 1993). However, there are two important advantages of the DRS representation that are important to the EA NLU approach. First, this representation allows for dynamic update of DRS because each variable and fact in a DRS is asserted independently. Adding or removing variables and facts from a DRS does not require retracting and editing a monolithic logical form. By design, this supports dynamic update of a discourse-level DRS as additional sentences are added to the discourse. Second, this representation allows back-chaining to individual facts in a DRS without regard for the nesting of that DRS. This is critical to enable the use of query-driven discourse-level processing in EA NLU, discussed in section 2.4.

DRT defines additional DRS embeddings corresponding to negation, implication and other quantifications. In EA NLU, these embeddings are used and expanded to represent the nested logical forms described in section 2.3.4. Implication is implemented in EA NLU using the predicate *implies-DrsDrs*, which takes two DRS (antecedent and consequent) as arguments. As in DRT, the antecedent DRS outscopes the consequent DRS such that the universe of the antecedent is valid in the consequent. Implication is used in this corpus of narratives primarily for numerical and qualitative quantification, as described in section 2.3.5. The example expression (35), repeated here:

(81) (thereExists group-of-cat1234
(and (isa group-of-cat1234 Set-Mathematical)
(cadinality group-of-cat1234 5)
(forAll cat1234 (implies (member cat1234 group-of-cat1234)
(isa cat1234 Cat)))))

translates to the DRS show in Figure 2.

Universe: group-of-cat1234

(isa group-of-cat1234 Set-Mathematical)
(cadinality group-of-cat1234 5)
(implies-DrsDrs DRS-antecedent DRS-consequent)

DRS-antecedent:

> Universe: cat1234
>
> (member cat1234 group-of-cat1234)

DRS-consequent:

> Universe:
>
> (isa cat1234 Cat)

Figure 2. DRS for the phrase "5 cats".

Modal operators such as *possible* and *willBe* are transformed in the same manner as the negation operator *not* in expression (75). This is not limited to a restricted set of operators; any higher-order predicate defined in the knowledge base may legally be used in a semantic frame that leads to nesting. The compositional semantics do not place any limit on nesting, it is determined entirely by the interaction of the frames and the grammar. Any time a quantified expression is nested as an argument in another expression, it can be transformed into a DRS. For example, the sentence:

(82) He wanted to move.

uses infinitive complement syntax as described above to generate a higher-order *desires* relation between the agent “he” and the clausal complement “to move”. Because the latter clause quantifies the movement event, it results in an embedded DRS structure show in Figure 3.

Universe: he1234

(desires he1234 (DrsCaseFn DRS-1234))

DRS-1234:

> Universe: move1234
>
> (isa move1234 Movement-TranslationEvent)
> (primaryObjectMoving move1234 he1234)

Figure 3. DRS for the sentence "He wanted to move."

2.3.14 Maintaining ambiguity with *choice sets*

The context-independent nature of compositional semantics is a key strength of this approach. It allows processing of complex, nested semantic forms without becoming computationally intractable. To enable this, ambiguities at the syntactic and semantic levels must be maintained in a way that is not combinatorially explosive. This is achieved at the syntactic level by supporting multiple parse trees, but constraining the supported syntactic patterns. At the semantic level, explicit *choice sets* are used. This packed representation was inspired by work with Xerox's XLE parser (Riezler, King, Crouch, & Zaenen, 2003).

A choice set is a logical form that expresses a semantic ambiguity. Ambiguities are introduced in a number of ways, beginning with the denotations and subcategorization frames in the knowledge base. This linguistic knowledge is set up within Cyc to be entered on a case-by-case basis – there is no systematic constraint that requires the set of cases to provide a certain completeness or consistency. This is important from a knowledge engineering standpoint, but places the burden on systems using the knowledge to deal with incompleteness and inconsistency. This is appropriate as a context-specific task has far more constraint to apply, and it somewhat mitigates the problem of trying to pre-engineer one perfect representation. The verb "throw", as an example, has eight frames in the current knowledge base. They include distinct word sense as well as differing syntactic cases of the same sense. Consider two of these frames:

(83) (verbSemTrans Throw-TheWord 1 TransitiveNPFrame
(and (isa :ACTION ThrowingAnObject)
(performedBy :ACTION :SUBJECT)
(objectActedOn :ACTION :OBJECT)))

(84) (verbSemTrans Throw-TheWord 1 DitransitiveNP-NPFrame
(and (isa :ACTION ThrowingAnObject)
(performedBy :ACTION :SUBJECT)
(objectActedOn :ACTION :OBJECT)
(toLocation :ACTION :INDIRECT-OBJECT)))

The simple sentence:

(85) I threw the ball.

generates a leaf constituent for the verb "threw". If the two frames above were the only semantic frames available for "threw" as a verb, then the semantics of the constituent would be the choice set expression of the form:

(86) (choiceSet (ChoiceSetFn FrameSemantics (ConstitFn verb1234 (SpanFn 1 2)))
(and (isa :ACTION ThrowingAnObject)
(performedBy :ACTION :SUBJECT)
(objectActedOn :ACTION :OBJECT))
(and (isa :ACTION ThrowingAnObject)
(performedBy :ACTION :SUBJECT)
(objectActedOn :ACTION :OBJECT)
(toLocation :ACTION :INDIRECT-OBJECT)))

Several logical functions are used in this expression to indicate the source of the choice.

(87) (SpanFn <start> <end>)

denotes the span in the tokenized input text beginning with the token at position <start> and ending before the position <end>. In the example it is the token “threw”.

(88) (ConstitFn <name> <span>)

denotes a named constituent generated by the parser over that span.

(89) (ChoiceSetFn <type> <arg1> ... <argn>)

denotes the reified choice set of type <*type*> with the specified type-specific arguments. In the case of a *FrameSemantics* choice set, the single argument is the leaf constituent that the choice set was generated from. The choice set connects this identifying information with a list of semantic translations, in this example the two available semantic frames. Because this ambiguity is transformed into an explicit expression, it is handled by the compositional semantics in exactly the same manner as an unambiguous expression. Both the role substitution and the semantic composition are unaffected by the ambiguity. This prevents combinatorial explosion while maintaining the context-independent nature of the process.

As part of the process of transforming the compositional semantics into the DRS representation, choice sets are extracted and reified in the working memory. This explicitly defines the disambiguation task at the sentence-level in terms of these choices sets. They are created for the frame semantics as described above, and also for parse trees, quantifier collisions and anaphoric references.

In the process of reification, invalid frames are discarded. Each semantic frame contains syntactic constraints that are used to aid disambiguation. Primarily, the subcategorization roles must be filled. If a certain frame expects a :OBJECT and there is none in the syntactic composition, then that frame can be discarded. Additionally, certain types of subcategorization frames place requirements on prepositional attachments. A *PPCompFrameFn* applies to a verb which requires a prepositional complement and specifies the particular preposition, as in:

(90) (verbSemTrans Carry-TheWord 2
(PPCompFrameFn TransitivePPFrameType In-TheWord)
(and (isa :ACTION Conveying-Stationary)
(conveyor-Stationary :ACTION :OBLIQUE-OBJECT)
(transportees :ACTION :OBJECT)))

This specifies that the :OBLIQUE-OBJECT role can only be filled from a PP that uses the term *in*, such as:

(91) He was carried in the stream.

Frames with unfilled roles may stem from this type of constraint as well as the lack of a binding to that role in the parse tree. In any case, such a frame is discarded.

2.3.15 Complexity in compositional frame semantics

Compositional frame semantics limits combinatorial complexity to two sources. The first is the grammar. The number of constituents for a given phrase within a sentence multiplies with the number of constituents anywhere else in the sentence that are not its descendants. In this practical approach, the use of controlled syntax limits this source of complexity. There are still a number of valid grammar-level ambiguities that reflect distinct semantic compositions in QRG-

CE, most notably propositional attachments and single/plural/generic nouns. The second source of complexity is nesting of higher-order predicates in the semantic translation. Every ambiguous nesting is a branch point that multiplies with every ambiguous nesting subordinate to it. Importantly, both unambiguous nestings and ambiguous facts that do not nest contribute only a constant factor to this problem. In this study, the number of semantic frames that present ambiguous nesting patterns is quite small. Further, the structural predicates such as modal operators and quantifiers are unambiguous in their nesting patterns, and they are exactly the ones most likely to nest with each other. Concepts such as desires, believes or intends are unlikely to nest more than two or three levels in a single sentence.

Because these sources of complexity are limited to well-understood patterns in the grammar and the composition, they are easy to monitor and appreciate in practice. While the grammar control is necessary to prevent intractability, no such restrictions have been necessary for the nesting of higher-order predicates. Over the set of narratives considered in this work, the compositional frame semantics process has performed consistently in the range of a few seconds per sentence. This is certainly not real-time performance, but practically sufficient for this work.

2.4 Query-driven discourse interpretation

The EA NLU approach is concerned with pragmatic language understanding, constrained by the concerns of a particular reasoning task over natural language. Cognitive models in particular provide a clear task with a well-defined model of the pragmatic concerns in performing that task. The goal of EA NLU is to facilitate natural language input to any cognitive model whose account of pragmatic understanding of natural language text can be formulated as a set of queries. Those queries are used to guide discourse-level interpretation.

The end result of discourse-level interpretation is a logical representation of the content of the discourse. This representation must be entailed by the source text and enable whatever reasoning is necessary for the cognitive model being considered. The interpretation is a serial process, where each subsequent sentence in the discourse updates that representation. As described in section 2.3, EA NLU uses compositional frame semantics to generate a DRS for each sentence. The discourse-level representation is also a DRS, the result of sequentially merging each sentence-level DRS into it. However, DRS merge is only the final step in the process for each sentence. First the system must address the issues of ambiguity and context. The sentence-level composition in EA NLU is able to operate under the assumption of context-independence only because disambiguation is being delayed. In addition, contextual reasoning is required to infer the implicit knowledge from the text that any non-trivial reasoning task will require. EA NLU relies on a model of pragmatic understanding accompanying the task, such as provided by a cognitive model in the form of a set of queries, to guide what inferences are made.

The context of interpretation for a sentence in a discourse consists of three elements. First, there is the discourse-level representation being constructed as a DRS. This DRS represents the described content of the discourse, that is, the world model that is being communicated. Second, there is the actual presentation of the discourse. This consists of the surface form of the sentences together with lexical/syntactic parsing information. Knowledge about how the discourse is presented, concepts such as word choice and sentence ordering, is not a part of the content of the discourse. Nevertheless it is a very significant part of the context within which a subsequent sentence is understood. This knowledge is stored in working memory in a separate microtheory, the *discourse case*. Third, there is the non-linguistic context, encompassing prior knowledge as well as the specific extra-linguistic circumstances of the communication. This broadly includes such things as cultural expectations, political climate, the mood of the hearer and the relationship between the communicators. A significant advantage to using cognitive modeling to scope understanding is that the experiments they are based on are specifically designed to control extraneous factors. The model itself bears the burden of accounting for all relevant context. Thus this third context contains the background knowledge and queries determined by the cognitive model.

Background knowledge is provided using the Cyc microtheory hierarchy. One or more microtheories, each of which possibly generalizes from a set of additional microtheories, are specified at the time of interpretation. Significantly, this background knowledge includes the complete set of horn-clause axioms used in the interpretation. The task model is defined by a set of facts, accessible from the specified microtheories, of the form:

(92) (queryForInterpretation <priority> <query-expression>)

The DRS translation of a new sentence in the discourse is interpreted in this context by attempting to prove each of these *query-expressions* by querying it. The expressions are first sorted by *priority*. One additional fact controls the process.

(93) (exhaustiveQueriesForInterpretation <microtheory>)

indicates that the set of expressions found in *microtheory* should be exhaustively queried until no new conclusions are proven. If this fact is not a part of the context, then each expression is queried only once. These two modes make the distinction between a task model that implements a hierarchical control structure and one that does not. The former does not generally require exhaustive proof, since the roots of the hierarchy are often designed to be independent of one another (otherwise they would not be separate roots). If this control is not encoded in the task model, then the more general approach of exhaustive proof is applicable.

As an example, the most general interpretation context is *EAGeneralQueriesMt*. This microtheory is intended to be included in the context of more specific task-models. It generalizes from two additional microtheories, *EAReferenceResolutionMt* and *EAStructuralRulesMt*. The former contains an axiomatization of general purpose reference resolution rules for pronouns and definite noun references. The latter contains additional axioms for reasoning across multiple contexts. There is only one query specified by this task model,

(94) (queryForInterpretation 0 (resolveUnresolvedReferences <sentence-identifier> ?var))

which invokes the most general purpose, knowledge poor rules for anaphora resolution. This results in attempts to prove the expression:

(95) (resolveReference
(ReferenceInSentenceFn <reference-var> <reference-sentence-identifier>)
(ReferentInDrsFn ?referent ?drs-identifier))

for each *reference-var* in the sentence-level DRS (which has the identifier *reference-sentence-identifier*) that does not already have such a fact asserted about it. The variables *?referent* and *?drs-identifier* uniquely identify a discourse variable already in the universe of a DRS within the discourse-level DRS. The axiomatization in *EAReferenceResolutionMt* is deliberately a last-effort heuristic approach that relies in part on language control. Resolution of pronouns and definite references is handled with a straightforward most-recent-candidate approach. Pronouns are restricted to noun-phrase entities (including gerund verbs), and pronoun reference is constrained by number and gender and prefers subjects. Definite references are resolved based on selectional restrictions from the Cyc ontology and can be used to refer to any entity. The pronouns “this” and “that” are restricted to refer to the preceding sentence head verb. More specific task models are expected to bring domain knowledge to bear on the task on anaphora resolution, which is necessarily ambiguous without it.

While a task model may contain any set of axioms, there are three general-purpose axiomatizations developed as part of EA NLU that have wide applicability. The first is the set of domain-general reference resolution heuristics. The other two deal with dialogue understanding and temporal ordering. As with *EAReferenceResolutionMt*, the goal is not to create complete axiomatizations, because that is not possible without context and intractable with it. Keeping with the practical principle, the goal is to create a set of axioms that are predictably limited when used independent of context. The expectation is that where more flexible reasoning is required, domain-specific, context-aware knowledge will be applied first.

2.4.1 Reasoning about temporal relations

In narrative text it is critical to understand the sequencing and overlaps of the situations and events being presented. I use Cyc’s representation of Allen’s *interval calculus* (J. F. Allen & Ferguson, 1994) for specifying temporal relations among reified instances of *Situation* (and thus *Event*). These relations provide considerable semantic breadth for a wide range of configurations among arbitrary numbers of *TemporalThings*, of which *Situation* is a specialization.

In order to infer temporal relations, this axiomatization adopts Reichenbach's *point of reference* in the representation of tense and aspect (Reichenbach, 1947). The point of reference combines with the point of speaking and the time of the event to fully represent the temporal aspects of a description of that event. In a sequence of event descriptions, temporal relations can be inferred in part based on continuity of the point of reference between subsequent descriptions. In this work I use the point of reference as a temporal cursor between sentences. Following the presentation of a sentence, the reference point can be said to be placed relative to some event in that sentence or a prior one. This is asserted in terms of two predicates, *placedAfter* and *placedDuring*, where the former indicates a narrative-specific immediacy – the specified event is the last relevant thing to happen prior to the cursor. The assertions take the form:

(96) *(placedAfter (ReferencePointFn <sentence identifier>) <event>)*

where the function *ReferencePointFn* returns the cursor following the specified sentence. When a new sentence is processed in this context, the position of the cursor from the prior sentence is used to evaluate the temporal positioning of each event in that sentence, as well as the positioning of the cursor following that sentence.

The simplifying assumption is made that the point of reference for a non-progressive verb is after rather than during the event interval. In general either could be the case, but given that the progressive aspect can be used to indicate a reference point during the interval, this makes a convenient point of language simplification. The result of this is that the simple narrative case – sequential, past-tense, non-progressive, imperfect verbs – is understood as sequential, non-overlapping events. Thus a discourse such as:

(97) I woke up. I brushed my teeth and ate breakfast.

results in a waking event followed by a brushing event followed by an eating event. This assumption is not made for stative verbs corresponding to *Situation* entities that are not instances of *Event* (e.g. "He is hot."). The point of reference for progressive verbs (situations as well as events) is assumed to be during the event or situation described, while the point of reference for perfect verbs can be safely assumed to be after the end of the situation or event.

Temporal relations for events in a new sentence can be inferred based on the position of the cursor following the prior sentence. Any imperfect, non-progressive, non-future event verb is assumed to describe an event that begins at the time of the prior cursor, which it then moves to its own reference point. Situations as well as progressive, perfect and future event verbs adopt the prior cursor as their reference point and do not move it.

Reasoning about implied temporal relations is more complex than these assumptions. There are cases where they do not hold, such as:

(98) I wrote a letter. I used a red pen.

which should not one event following another. There are also cases where more relations could be inferred, such as:

(99) He was holding a pen. He dropped it.

which implies that the holding ended at the dropping. It is also true that temporal reasoning extends beyond inferring the temporal semantics from presentation. Due to the binary nature of the interval calculus relations, there are many transitive axioms, upon which reasoning about the persistence and temporal overlap of fluents lies. In the general case, this explodes in the number of temporal entities, so those rules are not included in this axiomatization. Instead, this set of axioms provides a tractable, consistent baseline for general temporal reasoning. The burden for more powerful reasoning about temporal relations falls on task-specific axioms that can leverage contextual factors to constrain the problem.

2.4.2 Reasoning about dialogue

The first task in interpreting a dialogue is recognizing that a set of independent utterances are related to one another. This is approached in two ways here. First, it is possible that a narrative will explicitly assert that a dialogue is taking place, resulting in a *Conversation* event. If one or more *performedBy* roles for that event are also specified, then utterances attributed to those agents can be assumed part of the ongoing conversation (subject to temporal and spatial constraints).

As described in section 2.3.11, utterances are represented as communication events together with a nested description of the content of the communication. *Backward looking functions* (J. F. Allen & Core, 1997) are inferred over these communication events; specifically, the broad categories of *answer* and *elaboration*. The communication events in Cyc represent a *speech act* theory (Austin, 1975; Searle, 1969). These events, such as *Informing*, *RequestingInformation* and *ReplyingToAQuestion*, are identified through frame semantics (from terms such as "said", "asked" and "answered") along with the *senderOfInfo* and *recipientOfInfo* role relations. In cases where both roles are explicit across communication events, it is straightforward to identify turn taking and thus the backward looking functions. Similarly, successive utterances from one agent to another can be taken as elaboration, which allows inference of transitive responses.

That is, if a request for information is followed by an elaboration, a following response from the receiving agent is likely an answer to that request.

In most cases in natural language, the agent roles are not specified for each utterance. Rather, there is an assumption that subsequent utterances involve the same agents unless otherwise noted. Thus in the case:

(100) John asked Mary what time it was. She said she didn't know.

It must be assumed first that Mary is speaking to John in return and second that her utterance is an answer to his question. Similarly, if two agents are said to be having a conversation followed by one performing an utterance to an unspecified recipient, it can be assumed that the recipient is the other agent in the conversation. This is clearly an oversimplification as the content of the utterance may indicate a different recipient (or none at all), but once again that is a level of inference better suited to contextual, domain-specific reasoning.

2.4.3 Discourse merge

Once the task-specific queries have been completed for a sentence, the contents of the sentence-level DRS are merged into the discourse-level DRS. This process justifies universe variables and facts in the discourse-level DRS based on their counterparts in the sentence-level DRS, subject to two transformations. First, all sentence-level discourse variables are replaced by their referents if they have a *resolveReference* fact known about them. Second, if a fact of the form:

(101) (denotes <var1> <var2>)

is true in a sentence-level DRS than all instances of the discourse variable *var2* in that DRS (and any DRS scoped within that DRS) are replaced with the discourse variable *var1*. Thus recognition of intra-sentential reference is the task of proving facts of the form (101). QRG-CE recognizes some instances of this type of reference based on syntactic patterns.

2.5 Related work

2.5.1 In-depth semantic understanding

The challenge of generating formal representations of natural language narratives has been studied for some time. In early work it was largely assumed that such representations should be suitable for the kind of deep and broad reasoning we encounter in cognitive modeling tasks. Early work in narrative understanding (Charniak, 1972, 1977; Cullingford, 1978; Schank & Ableson, 1977) explored the role of non-linguistic world knowledge in the understanding process. By recognizing patterns of highly-structured world knowledge in narratives, these systems were able to use the expectation represented by unfilled roles in the knowledge to disambiguate and infer implicit facts. Wilensky (Wilensky, 1978) investigated the more specific and story-relevant case of knowledge about goals, plans and intentions. This line of work provided evidence not only for the role of world knowledge and expectation, but the power of these structured abstractions for explaining what makes a narrative coherent and controlling inferential complexity.

Later work in this same tradition done by (Lehnert, 1981) and (Dyer, 1983) turned to abstractions of storytelling to provide yet more explanatory power. Lehnert's *plot units* captured a certain predictable grammar of events in the plot of stories while Dyer's *thematic abstraction* *unit*s captured typical meanings in the form of plan failures. In both cases it was clear that such abstractions could provide higher level guidance and constraint to the understanding process. Further, recognition of these thematic concerns provides a deeper understanding of the meaning of a narrative.

This generation of work provides compelling evidence for explanation-based application of world knowledge to language understanding. It also demonstrates that prior knowledge of typical narrative constructs can be effectively leveraged to simplify understanding narrative accounts and deepen meaning. However, there were certain limitations. The difficulty of knowledge engineering was a major obstacle to scaling these systems, and their evaluations were limited to small examples. This eventually led to questions about whether there was a clear path to scaling (Lehnert, 1994). There was little follow-up work done in the larger research community, as attention shifted to less knowledge dependent formal logics and statistical processing methods. Subsequent work in meaningful semantic understanding of narratives turned to sophisticated logics and general-purpose inference mechanisms (Asher & Lascarides, 2003; Hobbs, Stickel, Appelt, & Martin, 1990; Ng & Mooney, 1990; Schubert & Hwang, 2000). However, the appeal to non-linguistic world knowledge became thinner and thinner. Simple examples used to prove formalisms are often taken without context and where necessary world knowledge is invoked there are deliberately vague assumptions about its availability.

2.5.2 Task model-driven understanding

Most current deep semantic understanding systems work within the practical constraint of a given task model or restricted domain. This approach has the dual benefit of limiting the breadth of background knowledge and working within a single, consistent deep model of the task at hand. My work could be viewed is an attempt to extend this category of systems by choosing a task – story understanding for cognitive modeling – that gains some of the benefits of considering distinct, well-defined tasks with far fewer limitations on the breadth of topics and modes of communication.

The *TRIPS* system (J. F. Allen et al., 2001) is the result of a long-term project in spoken dialogue systems aimed at human computer interaction modeled on human dialogue. TRIPS focuses on *practical dialogue*, defined as dialogue aimed at completing a concrete task. It follows the hypotheses that 1) practical dialogue competence is significantly simpler to achieve than general-purpose dialogue competence and 2) the bulk of the complexity in the language interpretation and dialogue management is independent of that task. This leads to a dual-layer task model, one for dialogue management and one for the concrete task. The dialogue management is built on an abstract problem-solving model that different tasks can then specialize. TRIPS uses a speech recognition module to identify words and input them to a best-first bottom-up chart parser also descended from Allen's textbook parser (J. F. Allen, 1994). A generic grammar and set of predicates is defined for practical dialogue and specialized by domain-specific lexical items and predicate mappings. The output of the parser is a sequence of speech acts which are understood in terms of the dialogue and task models. This approach heavily relies on expectation-based processing and strongly limits the breadth of possible semantic translations. By choosing that limitation, it is able to achieve inferential depth and robustness within its particular domain. The query-driven interpretation of EA NLU follows a similar approach, but does so in the context of much broader story understanding enabled by compositional frame semantics. The TRIPS practical dialogue model is ontologically rich and could certainly be axiomatized for reasoning in

EA NLU about practical dialogues (or practical dialogues in stories). Allen and his colleagues have demonstrated the breadth and robustness of the practical dialogue model by building several specialized systems on it, including the recent PLOW collaborative learning system (J. Allen et al., 2007).

2.5.3 Knowledge acquisition

Several research initiatives have approached the problem of acquiring formal knowledge from natural language text. Like this work, these systems face the common challenges of natural language understanding (e.g. syntactic ambiguity, word sense disambiguation, anaphora resolution, etc) in generating propositional content. However, the focus of these efforts has been on general facts and rules rather than episodic narratives. They attempt to extract sets of true facts from text in a way that is scalable to large corpora. By contrast, this work is concerned with representing the content of a narrative as a whole, requiring more complex relations between those true facts.

The Cycorp *TextLearner* project (Curits et al., 2006) leverages the knowledge in the Cyc knowledge base to generate an information-rich model of a document which can then be used as a guide for learning. TextLearner is concerned with learning both document contents and unfamiliar natural language constructs. To facilitate this, the system creates a detailed formal representation of the presentation aspects of the target text (i.e. tokenization, linkages, parse trees etc…) and defers resolution of ambiguities to broader reasoning. This is very similar to the way that my system divides compositional frame semantics from contextual reasoning. However, TextLearner chooses a different way of addressing the intractability of the general language understanding problem. Because the goal of the system is not to enable reasoning about a coherent account contained in a text, it pursues the more manageable goal of *semantic annotation*. That is, it identifies certain classes of facts that are contained in a text. This allows the system to avoid placing constraints on domain, vocabulary or grammar, which results in a higher degree of robustness and scalability. TextLearner also demonstrates the ability to reason about implicit content in the text which the authors refer to as *presuppositions*. This is limited by deliberate scoping to implications of anaphora such as the use of the pronoun “he” implying a gender classification of male for the referent. Much of the work on TextLearner is focused on learning rules for reading so as to improve its content extraction ability.

The *Boeing Language Understanding Engine (BLUE)* (P. Clark & Harrison, 2008) is intended to create general-purpose, formal representations of text content. BLUE was originally developed with the controlled language CPL (P. Clark et al., 2005), but has more recently been evaluated on unconstrained text. This is enabled by a clear distinction in pipeline stages from a broad-coverage parser and logic form generator to an initial logic generator and subsequent processing modules. Even where the latter modules are too narrow to cover the input, partial success can be attained based on the robust, shallow processing of earlier modules. The initial logic generator is based on simple syntactic rewrite rules, making it very similar to EA NLU’s composition process. However, BLUE does not benefit from extensive semantic frame knowledge enabling highly specific role relations and complex forms. Instead it relies on a small set of more general binary relations: *subject* (subject), *sobject* (syntactic object), *mod* (modifier), all the prepositions, *value* (for physical quantities), *number-of-elements* (for numbered plurals) and *named* (for proper names). This again allows more flexibility because the distinctions between correct and incorrect answers are fewer, but at the cost of semantic expressiveness. Subsequent processing in BLUE is separated into individual modules. Different strategies are used for word sense disambiguation, semantic role labeling, coreference, metonymy and *structural transformations*. The last module involves changes in the logical structure from the default, syntactically motivated form (e.g. verb reified as an event) to a desired semantic form (e.g. verb as an unreified relation). Excepting metonymy, EA NLU is able to address these phenomena based on the composition of semantic frames. Further, ambiguities are uniformly represented by choice sets and all forms of disambiguation rely on general query-driven back-chaining. As is often the case, this unification provides a consistent reasoning model at the expense of per-module engineering flexibility.

The *TextRunner* system (Banko, Cafarella, Soderland, Broadhead, & Etzioni, 2007) extracts relational tuples from very large scale corpora. It has been evaluated on a corpus of over 9 million web pages, demonstrating the ability to extract millions of triples with associated truth-probabilities representing both concrete facts and abstract assertions. TextRunner is entirely self-supervised, generating a classifier from a small corpus sample which is used in a single pass over the entire corpus to classify extracted tuples as to their trustworthiness. Finally, probabilities are assigned to each trustworthy tuple based on a probabilistic model of redundancy (Downey, Etzioni, & Soderland, 2005). This system demonstrates scalability in automated knowledge acquisition, but at the cost of limiting consideration to clusters of word triples.

2.5.4 Logics of language

Numerous logics have been proposed for representing natural language. Phenomena such as quantification, modality and non-monotonic update require significant increases in expressivity and complexity. In addition, narrative-specific logics have attempted to make explicit key constructs such as situations, actions, events and time. What is clear from this is that representing and reasoning with natural language requires an exceptionally high level of expressivity. What is less clear is how these logics should be compared against each other and, ultimately, against the task of narrative understanding. Each can demonstrate sufficiency and elegance in addressing particular linguistic phenomena and prove computational characteristics. However, I argue that the target of narrative understanding is too far off to reliably judge the merits of one approach vs. another. This view is supported by the continued prevalence of standard first-order logic (for well-understood complexity) and standard predicate calculus (for expressiveness) in implemented systems. That is not to suggest that these logics lack merit; only that it seems to me to be a premature optimization.

Episodic Logic (EL) (Schubert & Hwang, 2000) is a comprehensive framework that strives to "serve the full range of interpretive and inferential needs of general NLU." The goals of EL are to have natural language-like expressiveness, suitability for both world knowledge and semantic representation and straightforward derivability from surface forms. EL is specifically targeted towards representations of narrative, making episodes a core language feature. EL provides a more rigorous formalization of narrative-related linguistic phenomena such as adverbial modifiers, conditionals, actions and attitudes. This is similar in many ways to how my compositional frame semantics leverages higher-order expressions in CycL and DRT to model such phenomena. EL provides additional insights into the computational characteristics of these constructs, but sacrifices the flexibility and extensibility of a general higher-order predicate calculus combined with large-scale knowledge.

2.5.5 Statistical parsing

Statistical models of language and advances in wide-coverage grammars have made robust and efficient large-scale natural language processing a possibility (Kaplan et al., 2004; Matsuzaki, Miyao, & Tsujii, 2007). Such parsers can run over million and even billion-word corpora in reasonable amounts of time, identifying well-understood lexical-syntactic features. However, the translation from such surface features to deep semantic representations remains a problem even on a much smaller scale. Compositional frame semantics relies on semantic transformation rules that statistical approaches have not been able to model. EA NLU uses simplified English not because syntax is intractable on its own but because syntactic composition with deep semantics is. It is certainly the case that the current parsing strategy could be augmented by statistical techniques to gain more syntactic coverage. This would not increase semantic coverage, but it would provide a more graceful failure mode where semantic fragments are still positioned within a complete parse tree.

The Boxer (Bos, 2005) system, with the C&C Tools parser (Curran, Clark, & Bos, 2007), creates DRT-style representations over large-scale corpora. Boxer represents a significant push towards semantic depth in a statistical parsing system that maintains robustness at scale. Boxer predicates noun and verb entities using the root form of the lexical term (e.g. "dog" translates to *dog*(*x1*)). Adjectives and adverbs similarly predicate the modified noun or verb entity. These predicates are not grounded in an ontology (or otherwise axiomatized) for general reasoning. A standard set of event roles (e.g. agent, patient, etc) and prepositions introduce binary relations between entities. Boxer also handles the complex relations specified by DRT: negation, disjunction and implication. I view this as a complementary effort, given the similarities in the representational syntax. In fact, the output of Boxer can be viewed as an underspecified version of the output of EA NLU's compositional frame semantics. The predicates and role relations, defined in lexical-syntactic rather than semantic terms, are structured in similar DRS forms. As such, it seems likely that Boxer could provide useful heuristics in cases where EA NLU lacks grammatical or semantic coverage.

2.6 Conclusion

The practical EA NLU approach to language understanding facilitates the use of natural language input to cognitive modeling experiments by providing necessary semantic breadth while controlling computational complexity. Existing large-scale knowledge resources, ResearchCyc and COMLEX, provide lexical/semantic building blocks. Importantly, the subcategorization frames in Cyc express semantic translations in CycL, a sufficiently expressive logical language to capture higher-order relations. These building blocks are composed into complex, nested forms using compositional frame semantics. The interaction between the compositional process and the contents of the constituents is limited to syntactic roles and structural forms (quantifiers and modals), thus it is able to scale with the number of semantic frames in the knowledge base. Further, the composition maintains efficiency by being context-independent, such that general reasoning with arbitrary amounts of world knowledge is not necessary. This is made possible by the production of explicit choice sets in the semantic forms that maintain ambiguity for later reasoning. The output of the composition is transformed into nested DRS in a general-purpose reasoning environment. This enables both dynamic update and query-driven back-chaining with world knowledge in the discourse-level interpretation. EA NLU relies on the cognitive model itself to provide a specific reasoning task and pragmatic account of understanding for that task. Those pragmatic concerns are formulated as a set of queries that guide the discourse-level interpretation process. In the next chapter I will discuss how this novel integration facilitates natural language input to two specific cognitive models.

One of the key assumptions of this approach is that knowledge is necessary. This raises a question of scalability. Clearly the contents of the knowledge base do not cover all possible concepts or term semantics. This implementation provides three points of extension. For each point I am concerned with computational scalability as well as the level of expertise required to extend. First, lexical knowledge may be extended by adding entries to the COMLEX lexicon. Scaling with the number of entries is a simple database retrieval, and the knowledge base is built on a modern object-oriented database that is more than up to the task. Second, semantic breadth may be extended by adding subcategorization frames (or denotations) for terms. There is no real limit on the number of frames that could be created, making this a legitimate scaling concern. However, as stated above, the composition of the frames is impacted only by a set of syntactic roles and structural predicates. Those roles and structural predicates can increase as well, but practically the numbers are not significant. Expertise in knowledge representation is a very legitimate problem. By changing the representation task from the highly complex nested forms that a discourse might generate down to subcategorization frames, the expertise requirement is significantly reduced. Third, queries and axioms for different task models may be created. By using a query-driven interface, the system scales better to the task at hand rather than to the space of possible semantic interpretations for each sentence. It also allows multiple domain theories to be combined in the reasoning. However, there is a trade-off between the number of rules and facts involved and the cost of the reasoning. I will discuss this further in chapter 4. Aside from providing axioms for anaphora resolution, temporal reasoning and discourse reasoning, this approach does not reduce the expertise required for this third type of extension.

3.0 Natural Language Input for Cognitive Modeling

This chapter focuses on facilitating natural language input to three cognitive modeling experiments. The motivation for this is twofold: first, to provide a specific and measurable account of understanding narratives in support of the claim that EA NLU is an effective approach to that understanding. Second, to contribute both a theoretical approach and an implemented system that address the difficulty of formal representation and the problem of *tailorability.*

Cognitive modeling is based on the hypothesis that cognitive processes can be modeled as computation. By building a computational simulation of cognitive phenomenon, researchers are forced to make explicit the assumptions of psychological theories, often including unrecognized assumptions. This more rigorously defined model can then be validated against existing results, tested against new results and generate predictions suggesting future experiments. Typically, a cognitive simulation will use materials that are adapted from prior experiments with human subjects as input. Many of these input stimuli are narratives in natural language text. The model predicts responses to these stimuli, which constitute the output of the simulation. Cognitive modeling experiments provide a novel venue for natural language work. Each model provides a precise and evaluable case of understanding in terms of inferential capability. Taken across multiple models, the input texts are very broad in terms of the topics that they use and the kinds of reasoning tasks that participants are asked to carry out. Typically the representations used as input for the simulations are created by hand from the original texts, a process that is both labor-intensive and error prone. It also leads to the problem of *tailorability*, since the simulation authors (or people working closely with them) do the encoding of the formal representations. By automating the process of converting natural language to formal representation, or even semi-automating it, tailorability is reduced, and the plausibility of the simulation results is increased. EA NLU has been used in several cognitive modeling experiments including moral decision making (Dehghani, Tomai, Forbus, & Klenk, 2008), conceptual change (Friedman & Forbus, 2008) and blame attribution (Tomai & Forbus, 2008).

In this chapter I describe how EA NLU has been used in three cognitive modeling experiments. I provide evidence to support the claim that the practical language understanding approach implemented in EA NLU is an effective way to meet the inferential understanding requirements of cognitive models. I also discuss how this approach addresses the motivating problems of formal representation and reducing tailorability in cognitive simulation work. I start with a background discussion of two cognitive models that have used EA NLU in three simulation experiments. I then describe the experimental setup and detail how the semantic translation process performs in these experiments. This is followed by a brief discussion of the simulation results and how they pertain to the NL task. Finally, I contrast related approaches and conclude with general discussion.

3.1 Cognitive models

3.1.1 Moral decision making

While traditional models of decision-making in AI have focused on utilitarian theories, there is considerable psychological evidence that these theories fail to capture the full spectrum of human decision-making. In particular, research on moral reasoning has uncovered a conflict between utilitarian outcomes and normative judgments. Some researchers have proposed the existence of *sacred values* which evoke deontological moral rules (Baron & Spranca, 1997). Under those rules, the sacred values cannot be traded off, thus blocking utilitarian motives. Consider the starvation scenario from (Ritov & Baron, 1999) shown in Figure 4. The utilitarian decision in response to this scenario would send the convoy to the second camp, but participants tended to not divert the truck.

> A convoy of food trucks is on its way to a refugee camp during a famine in Africa. (Airplanes cannot be used.) You find that a second camp has even more refugees. If you tell the convoy to go to the second camp instead of the first, you will save 1000 people from death, but 100 people in the first camp will die as a result.
>
> Would you send the convoy to the second camp?

Figure 4: Moral decision making scenario from Ritov and Baron 1999

Given that life is a sacred value, people often refuse to take an action which would result in taking lives. (Tetlock, 2000) defines sacred values as "those values that a moral community treats as possessing transcendental significance that precludes comparisons, trade-offs, or indeed any mingling with secular values".

When sacred values are involved, people tend to be concerned with the nature of their action rather than the utility of the outcome. Baron and Spranca (1997) argue that people show lower *quantity sensitivity* to outcome utilities when dealing with sacred values. That is, they become less sensitive to the consequences of their choices, leading them to prefer inaction, even if it results in a lower outcome utility, over an action which violates a sacred value. The degree of outcome sensitivity varies with culture and the context of the scenario. (Lim & Baron, 1997) show that people in different cultures tend to protect different values and demonstrate different levels of sensitivity towards shared sacred values.

In addition to sacred values, the causal structure of the scenario affects people's decision-making. (Waldmann & Dieterich, 2007) show that people act more utilitarian, i.e., become more sensitive to action outcome utilities, if their action influences the agent of harm rather than the potential patient.

EA NLU was used in recent work with MoralDM (Dehghani et al., 2008), a cognitive model which captures these aspects of moral decision making. MoralDM incorporates two mutually exclusive modes of reasoning: utilitarian and deontological. If there are no sacred values involved in the case being analyzed, MoralDM applies traditional rules of utilitarian decision-making by choosing the action which provides the highest outcome utility. On the other hand, if MoralDM determines that there are sacred values involved, it operates in deontological mode and becomes less sensitive to the outcome utility of actions, preferring inactions to actions. For a given scenario a set of rules are applied to decide whether the case includes sacred values or not. An *orders of magnitude reasoning module* based on ROM(R) (Dauge, 1993) then calculates the relationship between the utility of each choice. Using the outcome of the orders of magnitude reasoning module, MoralDM utilizes a hybrid reasoning approach consisting of a *first-principles reasoning module* and an *analogical reasoning module* to arrive at a decision. The first-principles reasoning module suggests decisions based on rules of moral reasoning. The analogical reasoning module uses the *Structure Mapping Engine* (Falkenhainer, Forbus, & Gentner, 1989) to compare a given scenario with previously solved decision cases to determine whether sacred values exist in the new case and suggest a course of action. Using hybrid reasoning both gives the system the ability to tackle a boarder range of decision-making scenarios and provides a more cognitively plausible approach to decision-making.

3.1.2 Attribution of blame

There has been a significant amount of research in social psychology on how people infer internal states in others based on externally observable factors. A number of researchers, beginning with (Heider, 1983), have used *Attribution Theory* to investigate the conditions that will lead a perceiver to attribute some behavior, event or outcome to an internal disposition of the agent involved, as opposed to an environmental condition. Significantly, attribution is an entirely subjective process, based on the perceiver's understanding of the situation.

One significant area of this research deals with the attribution of moral responsibility and blame. When a human observes a set of actions and events leading to a negative outcome, there is a natural tendency to assign blame among the actors involved. Attribution of blame has been studied by (Shaver, 1985) and (Weiner, 1995). Shaver's theory posits five *dimensions of responsibility*: causality, intentionality, coercion, appreciation and foreknowledge. According to his theory, the process of attributing blame beings with assigning values along those dimensions based on observations surrounding a particular negative outcome. Those values are then used to indicate the relative amounts of responsibility, and ultimately blame, attributed to each agent. Responsibility here is "moral accountability", distinct from legal responsibility or the responsibilities of a formal office. Blame is moral condemnation that follows from responsibility for a morally reprehensible outcome.

For a given negative outcome, cause is defined as being an insufficient but necessary part of a condition which is itself unnecessary but sufficient for that result. Causal involvement in the negative outcome is a prerequisite for any responsibility to be assigned. Shaver characterizes intention as a scale of deliberateness with intentional at one end and involuntary at the other, such that the highest degree of intention should result in the strongest judgment of responsibility. Intention, however, can be moderated by coercion and appreciation. Coercion captures the force exerted by another agent which limits the available choices, from a social standpoint, for the agent in question. This could be through some direct threat or via an authority relationship. An agent who is coerced is assigned less responsibility than one who acts intentionally in the absence of coercion. Appreciation concerns the perceiver's judgment as to whether the agent in question has the capacity to understand that the outcome in question is morally wrong. If the agent does not have such capacity, they still bear some responsibility but are held exempt from blame. Foreknowledge is defined as the extent to which the agent was aware that an action would result in the outcome, prior to execution. Again, it is the perceiver's judgment of the knowledge the agent possessed that is evaluated. In the absence of intentionality, Shaver attributes responsibility based on foreknowledge.

In Shaver's model foreknowledge may be what the agent is thought to know (epistemic) or what the perceiver thinks the agent should have known (expected). However, it says little about the contribution of expected foreknowledge. This is not surprising as his model focuses on the perception of the agent's deliberative process. Weiner's model, by contrast, focuses on attribution of responsibility in cases of achievement and failure. In the case where an agent has failed to have expected foreknowledge, this model predicts that the perception of *causal controllability* over that failure determines the degree of responsibility attributed.

Mao (Mao, 2006) developed a computational model of responsibility assignment which models the judgments of *attribution variables* based on the dimensions of causality, intentionality, coercion and foreknowledge, and the attribution of blame following from those judgments. Mao presents an evaluation of her system against human data collected in a survey of 30 respondents. The survey presented four scenarios, variations of the well-known "company program" scenario used in (Knobe, 2003). The scenarios involve two agents, a chairman and a vice president, and a negative outcome of environmental harm. Figure 5 contains the text from one of the scenarios.

> The chairman of Beta Corporation is discussing a new program with the vice president of the corporation. The vice president says, "The new program will help us increase profits, but according to our investigation report, it will also harm the environment." The chairman answers, "I only want to make as much profit as I can. Start the new program!" The vice president says, "Ok," and executes the new program. The environment is harmed by the new program.

Figure 5: Corporate Program Scenario from Mao 2006

Each scenario was followed by a set of Yes/No questions intended to validate the judgments of intermediate variables, including the attribution variables, and a final question asking the respondent to score the blame each agent deserved on a scale of 1-6. Mao's system was evaluated by manually creating logical representations of the scenarios then running the inferential process to determine which of the agents in the scenario was to blame. The results of this evaluation are summarized in Table 1.

	Human Data		**Mao Model**		
	Chair	VP	Chair	VP	Degree
Scenario1	3.00	3.73		Y	Low
Scenario2	5.63	3.77	Y		Low
Scenario3	5.63	3.23	Y		Low
Scenario4	4.13	5.20		Y	High

Table 1: Results from Mao's evaluation of blame attribution

Mao's work is an important step towards modeling blame attribution. However, there are three limitations of interest here. First, as Mao observes, it uses Boolean values for attribution variables, whereas the underlying theories describe the dimensions of responsibility in terms of scalar values. Second, all blame is assigned to a single agent (or group of agents in a joint action). This is inconsistent with the human data in Mao's own experiment. Third, the degree of blame assigned by the system is limited to a value of *high* for intentional action and a value of *low* in the absence of intention. These assignments also do not match up with her data.

(Tomai & Forbus, 2008) presented a cognitive model of blame attribution using principles from *Qualitative Process Theory* (QP) (Forbus, 1984) that addresses these shortcomings. First, QP theory provides an appropriate level of representation for reasoning about social causality. Shaver and Weiner's theories, like many, are expressed in terms of continuous parameters such as "amount of intention" and "degree of foreknowledge". QP theory provides a principled way to formalize these ideas without over-constraining to numerical values. In particular, it rigorously defines ordinal relations between values and the inferential entailments that follow. This addresses the other two shortcomings in Mao's model by allowing the model to conclude relative amounts of blame between multiple agents. Blame is not limited to a binary distinction of Yes or No, nor arbitrary absolute values such as High and Low, but rather expressed in terms of agents being more or less to blame than other agents in similar circumstances.

3.2 Evaluation with cognitive models

EA NLU has been used as an integral part of three experiments using the two cognitive models described in section 3.1. In each experiment, the system enabled semi-automatic translation of text stimuli for use in the cognitive simulations. EA NLU was able to meet the requirements of semantic breadth posed by these simulations. Because the translation process is grounded in semantic frames in the knowledge base and uses the automated transformations described in chapter 2, it provides a more principled method for encoding the stimuli. For these experiments, a user-intervention mode was used to handle disambiguation of frames. This first step enabled evaluation of the content of the translation separately from evaluation of the disambiguation process. The following two chapters discuss automated disambiguation.

3.2.1 Experimental procedure

In all three of these experiments, the inputs are a set of stories told in natural language text. This text is manually translated to conform to the QRG-CE controlled language. This translation primarily entails altering the clausal structure of sentences – typically by dividing a long, complex sentence into several shorter ones. Each story is processed by EA NLU, resulting in a discourse-level DRS for that story. The compositional semantics builds the set of possible sentence-level DRS for each sentence and the associated choice sets defining the space of ambiguity. In these experiments, user intervention is used to disambiguate those choice sets (excepting anaphora resolution, which is handled automatically). Each cognitive model presents an account of understanding the narrative which can be formulated as a set of queries. This query model is used to drive discourse-level interpretation, as described in section 2.4, and the resulting discourse-level DRS is used as input to the cognitive model.

3.2.2 Understanding moral decision scenarios

MoralDM has been evaluated against studies by Ritov and Baron (Ritov & Baron, 1999) and Waldmann and Dieterich (Waldmann & Dieterich, 2007). In the experiment described here, four scenarios from each study were tested. The Ritov and Baron scenarios are shown in Table 2.

Original scenario	**QRG-CE translation**
A convoy of food trucks is on its way to a refugee camp during a famine in Africa. (Airplanes cannot be used.) You find that a second camp has even more refugees. If you tell the convoy to go to the second camp instead of the first, you will save 1000 people from death, but 100 people in the first camp will die as a result.	A convoy of trucks is transporting food to a refugee camp during a famine in Africa. 1000 people in a second refugee camp will die. You can save them by ordering the convoy to go to that refugee camp. The order would cause 100 people to die in the first refugee camp.
As a result of a dam on a river, 20 species of fish are threatened with extinction. By opening the dam for a month each year, you can save these species, but 2 species downstream will become extinct because of the changing water level.	Because of a dam on a river, 20 species of fish will be extinct. You can save them by opening the dam. The opening would cause 2 species of fish to be extinct.
Government support. Your office provides financial assistance to a plant employing 50 workers. If you withdraw this support (which will put 50 workers out of work) you can use the funds to support another plant, which employs 500 workers. Without government support, this second plant will close down.	Your office provides financial support to a plant that employs 50 workers. Another plant employs 500 workers. That plant will close. The closing will cause the 500 workers to be unemployed. You can save them by transferring the financial support. The transfer would cause the 50 workers to be unemployed.
Cutting forests. A logging company has the rights to 1000 square miles of old-growth forest. The company is willing to trade this land for 100 square miles of similar land, now part of a national park. You can give the smaller area to the company and make the larger area into a national park. The logging company will cut all the trees in whichever area it owns.	A logging company will deforest a 1000 square mile area. You can save it by trading a 100 square mile area. The trading will cause the logging company to deforest the 100 square mile area.

Table 2: Moral decision making scenarios from Ritov & Baron 1999

These were the first scenarios to be processed and evaluated using the current version of EA NLU. The simplified English was still being extended at that time, so the alterations here are more notable than for the stories in the later studies. In particular, a number of aspects of the story that do not impact the choice (as understood by MoralDM) were omitted at this point instead of being left in as noise for the reasoning system.

These scenarios each involve a situation where a default future is asserted followed by the possibility of intervention on the part of the participant, with associated consequences. In each scenario, choosing inaction (allowing the default future to come about) results in a higher utilitarian cost while acting results in a lower cost. They were designed exactly to test quantity sensitivity against a preference to avoid being a direct cause of a morally reprehensible outcome.

Consider the first scenario, regarding starvation in refugee camps. The first sentence:

(102) A convoy of trucks is transporting food to a refugee camp during a famine in Africa.

introduces four entities and two events. The convoy of trucks, the food and the refugee camp are indefinitely determined while the proper name Africa denotes the continent. The head verb of the sentence is the transporting event which is temporally positioned within a famine event.

Two parse trees are generated for this sentence using the rules of QRG-CE. The distinction between them is whether the prepositional phrase "in Africa" is attached to the verb "transporting" or the noun "famine". Four additional points of ambiguity arise from multiple subcategorization frames and denotations in the knowledge base. Table 3 shows the five sets of choices generated by EA NLU which were manually disambiguated in this experiment.

Source term (in bold)	Translations (selected option in bold)
convoy **of** trucks	a) **(possessiveRelation group-of-truck20258 convoy20231)** b) (generalizes convoy20231 group-of-truck20258)
trucks	a) (and (isa truck20258 (TransportViaFn Truck)) (transportees truck20258 convoy20231)) b) **(isa truck20258 Truck)** c) (isa truck20258 TruckDriver)
transporting	a) **(and (isa transport20294 TransportationEvent)** **(transportees transport20294 food20342)** **(transporter transport20294 convoy20231))** b) (isa transport20294 TransportationEvent) c) (isa transport20294 TransmittingSomething)
to a refugee camp	a) **(to-UnderspecifiedLocation transport20294 refugee-camp20515)** b) (to-Generic transport20294 refugee-camp20515) c) (to-UnderspecifiedLocation food20342 refugee-camp20515)
in Africa	a) **(in-UnderspecifiedContainer famine21557 (choiceForTerm africa))** b) (in-UnderspecifiedContainer transport20294 (choiceForTerm africa))

Table 3: frame semantics choice sets for "A convoy of trucks is transporting food to a refugee camp during a famine in Africa."

An important facet of frame semantics is that any given translation need be correct in only one possible case, meaning that there can be a great deal of ambiguity. It is also quite possible given the size of the knowledge base that certain frames might be invalid. These concerns motivate the abductive approach to disambiguation discussed in chapters 4 and 5.

One reason it is difficult to automatically disambiguate frames is that they tend to be *underspecified*. That is, the predicates used do not enforce strict argument type constraints nor are they rigorously axiomatized. Both predicates in the choices for the term “of”, *possessiveRelation* and *generalizes*, fall into this category. These types of predicates are useful for semantic translation because language is highly underspecified, but for the same reason it is difficult to use heuristics or simple constraints to identify invalid choices. One of the first tasks of reasoning over these representations is to specialize such predicates using contextual knowledge.

The compositional output for this sentence, based on these choices, is show in Figure 6.

Universe: group-of-truck20258 transport20294 convoy20231 famine21557
refugee-camp20515 food20342

(isa convoy20231 Convoy)
(possessiveRelation group-of-truck20258 convoy20231)
(isa group-of-truck20258 Set-Mathematical)

(implies-DrsDrs (DrsCaseFn DRS-3449947705-29007) (DrsCaseFn DRS-3449947705-29008))

DRS-3449947705-29007:

> Universe: truck20258
>
> (member truck20258 group-of-truck20258)

DRS-3449947705-29008:

> (isa truck20258 Truck)

(isa transport20294 TransportationEvent)
(transporter transport20294 convoy20231)
(transportees transport20294 food20342)
(to-UnderspecifiedLocation transport20294 refugee-camp20515)

(isa food20342 Food)
(isa refugee-camp20515 RefugeeCamp)

(isa famine21557 Famine)
(in-UnderspecifiedContainer famine21557 ContinentOfAfrica)
(temporallySubsumes famine21557 transport20294)

Figure 6: DRS for "A convoy of trucks is transporting food to a refugee camp during a famine in Africa."

This sentence demonstrates how typical event-entity types and relations are translated. In this sentence the composition is mostly quantification and conjunction; the exception being the plural "trucks" generating a universal implication. The remaining three sentences in this scenario present more complex ideas. They establish a default future where 1000 people in a second camp die, a possible intervention to save those people and the consequence of abandoning the

100 people in the first camp. The contrast between the default future and the possibility of intervention requires the semantic breadth to distinguish between what has already occurred and two distinct possible futures. Additionally, proper numerical quantification of the sets of people is an important part of understanding the choice being presented. These sentences also present higher order relations based on ordering a certain action and causing an outcome.

The second sentence:

(103) 1000 people in a second refugee camp will die.

uses the modal operator *willBe* to nest the hypothetical future where the 1000 people die as shown in Figure 7.

```
Universe: group-of-people2826

(cardinality group-of-people2826 1000)
(isa group-of-people2826 Set-Mathematical)

(willBe (DrsCaseFn DRS-3456227871-5344))

DRS-3456227871-5344:
  Universe: die3835

  (isa die3835 Dying)
  (objectOfStateChange die3835 group-of-people2826)

(implies-DrsDrs (DrsCaseFn DRS-3456227871-5342)
                (DrsCaseFn DRS-3456227871-5343))

DRS-3456227871-5342:
  Universe: people2826

  (member people2826 group-of-people2826)

DRS-3456227871-5343:
  Universe: refugee-camp3092 SERIES2920

  (in-UnderspecifiedContainer people2826 refugee-camp3092)
  (isa refugee-camp3092 RefugeeCamp)
  (isa people2826 Person)
  (nthInSeries refugee-camp3092 SERIES2920 2)
```

Figure 7: DRS for "1000 people in a second refugee camp will die."

The third sentence:

(104) You can save them by ordering the convoy to go to that refugee camp.

uses the modal operator *possible* to embed the hypothetical future where the order is given. The order event accepts an infinitive complement to express the content of that order, resulting in another nested DRS. This is shown in Figure 8.

```
Universe: you5350

(possible (DrsCaseFn DRS-3456230719-8304))

DRS-3456230719-8304:
    Universe: order5456 them5421 save5394 convoy5561

    (isa save5394 RescuingSomeone)
    (performedBy save5394 you5350)
    (beneficiary save5394 them5421)

    (causes-EventEvent order5456 save5394)

    (isa order5456 Ordering-CommunicationAct)
    (performedBy order5456 you5350)
    (recipientOfInfo order5456 convoy5561)
    (infoTransferred order5456 (DrsCaseFn DRS-3447201937-128237))

    DRS-3447201937-128237:
        Universe: refugee-camp6829 go5652

        (isa go5652 Movement-TranslationEvent)
        (primaryObjectMoving go5652 convoy5561)
        (toLocation go5652 refugee-camp6829)))
        (isa refugee-camp6829 RefugeeCamp)
        (isa convoy5561  Convoy)
```

Figure 8: DRS for "You can save them by ordering the convoy to go to that refugee camp."

The final sentence in this scenario:

(105) The order would cause 100 people to die in the first refugee camp.

expresses a causal statement combining with the modality of the term "would" and numerical quantification. This results in the nested DRS shown in Figure 9.

```
Universe: order29022

(possible-Historical (DrsCaseFn DRS-3449953357-29568))
(isa order29022 Ordering-CommunicationAct)

DRS-3449953357-29568:
    Universe: cause29074

    (causes-ThingProp order29022 (DrsCaseFn DRS-3449953357-29569))

    DRS-3449953357-29569:
        Universe: refugee-camp29362 group-of-people29113 die29166 SERIES29288

        (isa group-of-people29113 Set-Mathematical)
        (cardinality group-of-people29113 100)
        (implies-DrsDrs (DrsCaseFn DRS-3449953357-29570) (DrsCaseFn DRS-3449953357-29571))

        DRS-3449953357-29570:
            Universe: people29113

            (member people29113 group-of-people29113)
        DRS-3449953357-29571:

            (isa people29113 Person)


        (in-UnderspecifiedContainer group-of-people29113 refugee-camp29362)
        (isa refugee-camp29362 RefugeeCamp)
        (nthInSeries refugee-camp29362 SERIES29288 1)

        (isa die29166 Dying)
        (objectOfStateChange die29166 group-of-people29113)
```

Figure 9: DRS for "The order would cause 100 people to die in the first refugee camp."

These sentence-level DRS are interpreted as a discourse by EA NLU. The discourse-level processing relies on the pragmatics of understanding defined by the MoralDM first-principles reasoning module. This account of understanding identifies choices and their consequences. It is formulated as queries of the form:

(106) (isa ?selecting SelectingSomething)

(107) (choices ?selecting ?action)

(108) (causes-PropSit (chosenItem ?selecting ?action) ?outcome-event)

The Cyc collection *SelectingSomething* represents all events where a choice is made, with the role relation *choices* specifying to the known options and the role relation *chosenItem* specifying the option selected. Query (108) proves the primary outcomes explicitly stated with the choice. In the example scenario, the representation of the choice in the discourse is the set of facts:

(109) (isa sel131949 SelectingSomething)

(110) (choices sel131949 order131049)

(111) (choices sel131949 inaction131950)

(112) (causes-PropSit (chosenItem sel131949 inaction131950) die128829)

(113) (causes-PropSit (chosenItem sel131949 order131049) save128937)

There are two key bridging inferences that allow these queries to be proven. First, the presence of a default future (where 1000 people die) combined with an opportunity for intervention on behalf of the choosing agent (the listener) implies that inaction is a choice. Second, the abstraction of *saving* drives inferential attention to events in the description that the beneficiary may be being saved from. The expected future modality of the first set of deaths makes it a reasonable candidate. The saving event is therefore inferred to also bring about the prevention of those deaths, as expressed by the set of facts:

(114) (causes-SitSit save128937 prevent131948)

(115) (isa prevent131948 PreventingSomething)

(116) (performedBy prevent131948 you128898)

(117) (preventedSit prevent131948 die128829)

Combined with the outcome events and numerical quantification as expressed in Figure 9, these facts capture the nature of the trade-off being presented in this scenario. They are proved as part of the discourse-level interpretation, along with anaphora resolution, to create a discourse-level DRS that captures the entities and events of the scenario as well as the abstractions of choice.

The Waldmann and Dieterich scenarios each present a similar trade-off situation, but differ in that each scenario consists of a pair of scenarios designed to capture the distinction between agent and patient intervention. These are shown in Table 4.

Original scenario	**QRG-CE translation**
In a restaurant, a bomb threatens to kill 9 guests. The bomb could be thrown onto the patio, where 1 guest would be killed. In a restaurant, a bomb threatens to kill 9 guests. One guest could be thrown on the bomb, which would kill this 1 guest.	A bomb in a restaurant will kill 9 people. You can save them by throwing the bomb onto the patio. The throwing will kill 1 person. A bomb in a restaurant will kill 9 people. You can save them by throwing 1 person on the bomb. The throwing will kill that person.
A high-speed train is about to hit a bus with 10 passengers. An employee of the train company could redirect the train to a side track where a bus with 2 passengers is sitting. A high-speed train is about to hit a bus with 10 passengers. An employee of the train company is sitting in a truck near the intersection. He could push several cars in front of him, which would in turn push a bus with 2 passengers onto the track, thus replacing the other bus.	A high-speed train will hit a bus with 10 passengers. You can save them by redirecting the train to a side track. A bus with 2 passengers is on the side track. The redirecting will cause the train to hit that bus. A high-speed train will hit a bus with 10 passengers. You can save them by pushing a bus with two passengers. The pushing will cause the train to hit that bus.
A virus causing paraplegia threatens 4 patients. Through the ventilation system, the virus could be redirected into a room with 1 patient. A virus causing paraplegia threatens 4 patients. The bone marrow of 1 patient could save them. However, the required procedure would lead to paraplegia in this patient.	A virus will cause 4 patients to have paraplegia. You can save them by redirecting the virus into a room with 1 patient. The redirecting will cause that patient to have paraplegia. A virus will cause 4 patients to have paraplegia. You can save them by obtaining bone marrow from 1 patient. The obtaining will cause that patient to have paraplegia.
A torpedo threatens a boat with 6 soldiers. Destroying the torpedo by remote control would sink a nearby submarine with 3 soldiers. A torpedo threatens a boat with 6 soldiers. Three soldiers could be ordered to move their boat in a way that would divert the torpedo from the original target to their boat.	A torpedo will sink a boat with 6 soldiers. You can save them by destroying the torpedo by remote control. The destroying will sink a nearby submarine with 3 soldiers. A torpedo will sink a boat with 6 soldiers. You can save them by ordering a boat with 3 soldiers to intercept it. The order will sink the boat with 3 soldiers.

Table 4: Moral decision making scenarios from Waldmann and Dietrich 2007

These scenarios present a possible intervention, and thus a choice, in the same manner as the Ritov and Baron scenarios. Numerical quantification, modality and higher-order expressions such as cause are handled in the same way. The differences in entity and event types and roles are reflected in different semantic frames, but do not require a different composition. Within the proposed interventions for each scenario, a distinction is made as to the object of the intervening action. For example, consider the bomb in the restaurant scenarios. In the first version of the pair, the subject is being offered the possibility of acting on the bomb, which is the agent of harm. In the second version, the subject is being offered the possibility of acting on a person who would then be patient of harm (the one killed). This distinction is captured by the role relation *objectActedOn* between the throwing event in each version and the bomb or the person. That event must also be understood to cause the killing event in the next sentence which in turn has its own *organismKilled* role relation indicating the patient. The MoralDM task model is able to query for such relations to see, in this case, if the *objectActedOn* matches the *organismKilled*, a case of patient intervention.

The first-principles reasoning module in MoralDM must be able to prove all these aspects of understanding for a scenario in order to make a decision. The specific events and their relations inform the presence of culture-specific sacred values that invoke MoralDM's deontological reasoning mode. Causal structure, particularly focusing on the choice being presented and its consequences, as well as quantification and role relations provide the antecedents for either a utilitarian or a deontological choice. In all eight cases in this experiment, EA NLU was able to generate a representation that was sufficient to the task and enabled the first-principles reasoning module to produce a decision that matched the human data. Figure 10 shows the complete query model for MoralDM. The macro predicate *lookupOnly* prevents back-chaining and is used in the priority 1 queries to retrieve bindings from the priority 0 query without redundant effort.

```
(queryForInterpretation 0 (and (isa ?selecting SelectingSomething)
                               (choices ?selecting ?action)
                               (causes-PropSit (chosenItem ?selecting ?action) ?outcome-event))

(queryForInterpretation 1
        (and (lookupOnly (causes-PropSit (chosenItem ?selecting ?action) ?outcome-event))
             (violationOfSacredValue ?outcome-event))

(queryForInterpretation 1 (and (lookupOnly (choices ?selecting ?action))
                               (isa ?action Inaction))

(queryForInterpretation 1
        (and (lookupOnly (choices ?selecting ?action))
             (lookupOnly (causes-PropSit (chosenItem ?selecting ?action) ?outcome-event)))
             (objectActedOn ?action ?patient)
             (objectActedOn ?outcome-event ?patient)))

(queryForInterpretation 1
        (and (lookupOnly (choices ?selecting ?action))
             (lookupOnly (causes-PropSit (chosenItem ?selecting ?action) ?outcome-event))
             (objectActedOn ?action ?agent)
             (doneBy ?outcome-event ?agent)))
```

Figure 10: Query model for MoralDM

Across the eight moral decision making scenarios, the discourse-level representation constructed by EA NLU included an average of 66 facts, 50 of which came directly from the compositional semantics and 16 of which were inferred during the discourse interpretation. Of those 66 facts, an average of 46, or 69%, were used by the MoralDM first-principles reasoning module to reach a correct decision.

The analogical reasoning module in MoralDM also uses these same representations to suggest decisions based on similarity with prior, correctly answered scenarios. In the first of the eight cases, regardless of order, this module could not return an answer because its library of prior cases was empty. Over the remaining seven cases, the analogical reasoning module was able to use the representations generated by EA NLU to make the correct decision in five of them. In the remaining two cases, the retrieval of a similar scenario failed (the module found no sufficiently similar prior scenario) thus returning no answer. Overall, MoralDM made the correct choice in all of the scenarios ($p < 0.005$).

3.2.3 Understanding blame attribution scenarios

In this experiment, EA NLU was used to process the “corporate program” scenarios from (Mao, 2006) for input to my cognitive model of blame attribution. The scenarios and their QRG-CE translations are show in Table 5.

Original scenario	QRG-CE translation
The vice president of Beta Corporation goes to the chairman of the board and requests, "Can we start a new program?" The vice president continues, "The new program will help us increase profits, and according to our investigation report, it has no harm to the environment." The chairman answers, "Very well." The vice president executes the new program. However, the environment is harmed by the new program.	The vice president of a corporation asks the chairman of the corporation, "Can we start a new program?" The vice president says, "The new program will increase profits. It will not harm the environment." The chairman answers, "Very well." The vice president executes the new program. However, the new program harms the environment.
The chairman of Beta Corporation is discussing a new program with the vice president of the corporation. The vice president says, "The new program will help us increase profits, but according to our investigation report, it will also harm the environment." The chairman answers, "I only want to make as much profit as I can. Start the new program!" The vice president says, "Ok," and executes the new program. The environment is harmed by the new program.	The chairman of a corporation is discussing a new program with the vice president of the corporation. The vice president says, "The new program will increase profits. Also, it will harm the environment." The chairman answers, "I only want to increase profits. Start the new program!" The vice president says, "Okay." The vice president executes the new program. The new program harms the environment.
The chairman of Beta Corporation is discussing a new program with the vice president of the corporation. The vice president says, "The new program will help us increase profits, but according to our investigation report, it will also harm the environment. Instead, we should run an alternative program, that will gain us fewer profits than this new program, but it has no harm to the environment." The chairman answers, "I only want to make as much profit as I can. Start the new program!" The vice president says, "Ok," and executes the new program. The environment is harmed by the new program.	The chairman of a corporation is discussing a new program with the vice president of the corporation. The vice president says, "The new program will increase profits. Also, it will harm the environment." The vice president says, "Instead, we should execute an alternative program. It will increase profits less than the new program, but it will not harm the environment." The chairman answers, "I only want to increase profits. Start the new program!" The vice president says, "Okay." The vice president executes the new program. The new program harms the environment.
The chairman of Beta Corporation is discussing a new program with the vice president of the corporation. The vice president says, "There are two ways to run this new program, a simple way and a complex way. Both will equally help us increase profits, but according to our investigation report, the simple way will also harm the environment." The chairman answers, "I only want to make as much profit as I can. Start the new program either way!" The vice president says, "Ok," and chooses the simple way to execute the new program. The environment is harmed.	The chairman of a corporation is discussing a new program with the vice president of the corporation. The vice president says, "There is a simple way to execute the program and a complex way to execute the program. Both ways will increase profits equally. Also, the simple way will harm the environment." The chairman answers, "I only want to increase profits. Start the new program either way!" The vice president says, "Okay." The vice president chooses the simple way to execute the new program. The new program harms the environment.

Table 5: Corporate program scenarios from Mao 2006

Given the four scenarios, there are eight agents (a chairman and a vice president in each) and thus twenty-eight possible comparisons between two agents. Where Mao's evaluation inferred a single agent to blame in each scenario, this evaluation considered the entire set of comparisons. I claim that this is a more informative experiment, and that it better fits the survey data collected by Mao. The participants in this survey were not asked which agent was to blame in each scenario, or even which agent was more to blame. Rather, they were presented with the four scenarios one at a time and asked after each one to assign a numerical amount of blame to each of the two agents individually using the same 1-6 scale. Given the deliberate similarity of the scenarios, the participants could not help but score agents relative to the scores they gave agents in prior scenarios. Given the random ordering of the surveys, the aggregate results present a relative scale among the eight agents in the four scenarios. Thus, ordinal relations among all pairings of the agents are a more informative measure for a cognitive model over this data.

The Mao scenarios each involve a conversation between the two agents where information regarding foreknowledge, intention and coercion relative to a possible future action is discussed. This conversation results in execution of the possible action which in turn causes environmental harm. Understanding this scenario, from the point of view of this cognitive model, involves being able to turn observations of external behavior into judgments along Shaver's dimensions of responsibility. The model represents attribution variables for intentionality, coercion, foreknowledge and responsibility as nonnegative continuous parameters. For example, the amount of foreknowledge an agent has that an action was going to cause an outcome is represented with the functional term:

(118) (ForeknowledgeFn <agent> (causes-SitSit <action> <outcome>))

Making an attribution means assigning either a qualitative value or an ordinal relation to such a term with respect to a particular time interval. Qualitative values are expressed as *limit points* corresponding to changes in qualitative state. For these parameters, valid limit points are *None* and *Certain* with the value *Some* representing the interval between them. Continuing with the example of causal foreknowledge, the representation of an attribution of a qualitative value is queried as:

(119) (attributedValue Foreknowledge
(ValueDuringFn (ForeknowledgeFn ?agent (causes-SitSit ?action ?outcome))
?time-interval)
?qualitative-value)

The representation of an ordinal relation is similarly queried as:

(120) (greaterThan
(ValueDuringFn (ForeknowledgeFn ?agent1 (causes-SitSit ?action1 ?outcome1))
?time-interval1)
(ValueDuringFn (ForeknowledgeFn ?agent2 (causes-SitSit ?action2 ?outcome2))
?time-interval2))

Judgments of causality remain Boolean, as that is the extent of their impact in Shaver's model. Causal contribution is inferred for an agent who performs an action which causes an outcome or an agent who coerces a performing agent. The former situation is understood from the text based on a *performedBy* role relation to an event that *causes* the outcome in question. This is expressed in a straightforward manner by the concluding pair of sentences in each scenario, such as from the second scenario:

(121) The vice president executes the new program. The new program harms the environment.

The latter situation depends on coercion, explained in detail below.

These scenarios are entirely dialogue-based, save the final action and outcome, so the understanding of foreknowledge, intention and coercion comes from what is said. Explicit communication of an expected future outcome entails attribution of some amount of foreknowledge of that outcome to both the speaker and the hearer. This foreknowledge may be accurate or not. Unless the speaker qualifies the expected outcome in some way, the system infers equality to an upper limit point of certainty. It does not address the issue of deception. For example, in the second scenario, the vice president clearly states the outcome of environmental harm resulting from the new program:

(122) The vice president says, "The new program will increase profits. Also, it will harm the environment."

The DRS for this sentence is shown in Figure 11.

```
Universe: vice-president29698 say29706

(isa say29706 Informing)
(senderOfInfo say29706 vice-president29698)
(isa vice-president29698 VicePresidentOfOrganization)

(infoTransferred say29706 (and (DrsCaseFn DRS-3450011203-30210)
                               (DrsCaseFn DRS-3450011203-30211)))

DRS-3450011203-30210:
  Universe: program29782

  (isa program29782 NewArtifact)
  (isa program29782 Project)

  (willBe (DrsCaseFn DRS-3450011203-30212))

  DRS-3450011203-30212:
    Universe: group-of-profit29883 increase29843

    (isa increase29843 IncreaseEvent)
    (doneBy increase29843 program29782)
    (objectActedOn increase29843 group-of-profit29883)

    (isa group-of-profit29883 Set-Mathematical)
    (implies-DrsDrs …)

DRS-3450011203-30211:
  Universe: it29576

  (willBe (DrsCaseFn DRS-3450010554-29678))

  DRS-3450010554-29678:

    Universe: environment29648 harm29617

    (isa environment29648 EcologicalRegion)
    (isa harm29617HarmingSomething)
    (doneBy harm29617it29576)
    (thingHarmed harm29617environment29648)
```

Figure 11: DRS for stating environmental harm

The abbreviated *implies-DrsDrs* in Figure 11 is the plural implication that the members of the set *group-of-profit29883* are of type *Profits*. This is in fact an error, instantiating a plural for this mass noun. However, it cannot be corrected without world knowledge about profits, a step that is left to the task-driven discourse-level processing. If the task does not require domain reasoning about the nature of profits, which this one does not, then the error will be ignored.

Based on the representation of this sentence, and the dialogue inference that the chairman was the recipient of the utterance (based on the statement that they are discussing), both agents are attributed certain foreknowledge regarding the environmental harm at the time of this utterance. The temporal interval over which that attribution holds is assumed to begin before the utterance for the vice president and during the utterance for the chairman. No assumption is made as to when it ends. This results in the assertion (for the vice president):

```
(123)  (attributedValue Foreknowledge
          (ValueDuringFn (ForeknowledgeFn vice-president29541
                                                     (causes-SitSit program29503 harm29617))
                         interval29620)
          Some)
```

where *interval29620* is a dynamically created variable for the time interval over which the attribution is believed to hold. This attribution contrasts with the first scenario where the vice president and chairman lack foreknowledge of the environmental harm as captured by the sentence:

(124) *The vice president says, "The new program will increase profits. It will not harm the environment."*

In that scenario the query for attributed foreknowledge fails.

A significant distinction is made between evidence of act and outcome intention, following (Weiner, 2001). It is assumed that an agent intends any action that they perform or order performed. If the action is known by the agent to have only one outcome, then that outcome is also intended. There is considerable philosophical discussion on whether foreknowledge of multiple outcomes implies intention of all those outcomes. Shaver claims a judgment of intention presupposes epistemic foreknowledge, but not the other way around (Shaver, 1985). Conversely, Bratman argues that epistemic foreknowledge combined with action must imply intention (Bratman, 1990). Acknowledging these different positions, the system makes the weaker inference that when an agent is certain of an outcome and performs or authorizes the action, it implies only some non-zero level of intention. When an agent orders an action that has multiple alternative outcomes and the performing agent is allowed to choose between them, outcome intention is entailed only for the performing agent. In the second scenario, both agents are attributed foreknowledge at the time of sentence (122). If that foreknowledge is understood to persist through the order given by the chairman, he will also be attributed some intention. If it persists through the actual execution of the program, the vice president will be attributed likewise. Attributions are assumed to persist until they meet an event that provides evidence for the attribution of a different value along the same dimension. In the second scenario, there is never any indication that the foreknowledge of environmental harm is called into question. The resulting assertion (for the vice president) is:

(125) (attributedValue Intention
(ValueDuringFn (IntentionFn vice-president29541 program29503 harm29617)
interval29632)
Some)

where *interval29620* is again a dynamically created variable for the time interval over which the attribution is believed to hold. This interval is to have an *overlaps* relation with the execution of the new program. The other three scenarios are designed to imply different attributions of intention. In the first scenario, there is no foreknowledge, so the attribution query fails. In the third scenario the vice president shows a clear lack of prior intention based on his recommendation that they choose a different program that will not harm the environment. He is therefore attributed a lack of intention prior to coercion by the chairman and attribution of some intention after. This is used in later reasoning in the model to assume that the coercion was stronger than in a case where clear lack of intention is not present. The chairman in the fourth scenario is attributed no intention based on his abdication of responsibility.

Coercion is applied by one agent to another via an action with regard to a particular action and outcome to bring about. The amount of this coercion is represented by a term of the form:

(126) (CoercionFn <coercer> <coerced> <coercion event> <action> <outcome>)

The distinction between action and outcome intent applies to coercion as well. Where an imperative command to act is given by an agent in a position of authority, some amount of action coercion is inferred. It may or may not be effective – this is known only by comparison with later actions. The authority of the chairman in these scenarios is inferred by world knowledge about the relationship between chairpersons and vice presidents within corporate structure combined with the imperative command utterance given by the chairman.

Outcome coercion is inferred by the same logic as outcome intention: both agents must have foreknowledge of the outcome at the time of the coercion and other outcome options must not be equally available to the coerced agent. An additional complication arises in that an agent with prior intention is not coerced by being ordered to do what he or she already intended. This necessitates again temporal comparisons as to the temporal relations between the interval over which intention is attributed and events in the scenario. In the second scenario, the chairman is applying coercion to the vice president, beginning at the time of the utterance:

(127) The chairman answers, "I only want to increase profits. Start the new program!"

The resulting attribution is asserted as:

```
(128)  (attributedValue Coercion
          (ValueDuringFn
             (CoercionFn chairman29512 vice-president29541
                        answer30123 program29503 harm29617)
             interval30992)
          Some)
```

where *answer30123* is the utterance event in sentence (127) and *interval30992* is a time interval that *startsDuring* that utterance and is ultimately found to have an *overlaps* the execution of the new program. Again, the variations in the scenarios imply different instances of coercion. In the first scenario, the lack of foreknowledge causes the query for attribution of coercion to fail. Coercion is attributed in the third scenario as in the second. In the fourth scenario the query again fails due to the freedom of choice allowed the vice president.

The complete query model for this model of blame attribution is shown in Figure 12.

```
(queryForInterpretation 0 (responsibleByActionFor ?agent ?action ?outcome))

(queryForInterpretation 0 (responsibleByCoercionFor ?agent ?coercion-event ?action ?outcome))

(queryForInterpretation 1
        (and (lookupOnly (responsibleByActionFor ?agent ?action ?outcome))
             (attributedValue
                 Foreknowledge
                 (ValueDuringFn  (ForeknowledgeFn ?agent (causes-SitSit ?action ?outcome)) ?action)
                 ?qualitative-value)))

(queryForInterpretation 1
        (and (lookupOnly (responsibleByCoercionFor ?agent ?coercion-event ?action ?outcome))
             (attributedValue
                 Foreknowledge
                 (ValueDuringFn  (ForeknowledgeFn ?agent (causes-SitSit ?action ?outcome)) ?action)
                 ?qualitative-value)))

(queryForInterpretation 1
        (and (lookupOnly (responsibleByActionFor ?agent ?action ?outcome))
             (attributedValue
                 Intention
                 (ValueDuringFn (IntentionFn ?agent ?action ?outcome) ?interval)
                 ?qualitative-value)))

(queryForInterpretation 1
        (and (lookupOnly (responsibleByCoercionFor ?agent ?coercion-event ?action ?outcome))
             (attributedValue
                 Intention
                 (ValueDuringFn (IntentionFn ?agent ?action ?outcome) ?interval)
                 ?qualitative-value)))

(queryForInterpretation 1
        (and (lookupOnly (responsibleByActionFor ?coerced ?action ?outcome))
             (attributedValue
                 Coercion
                 (ValueDuringFn (CoercionFn ?coercer ?coerced ?coercion-event ?action ?outcome) ?interval)
                 ?qualitative-value)))
```

Figure 12: Query model for blame attribution

Additional reasoning about temporal relations between intervals and ordinal relations between attributed values is done by the model of blame attribution once the complete discourse-level DRS has been constructed.

EA NLU was used to process the four Mao scenarios, resulting in discourse-level representations of the narratives and the attributions of causality, foreknowledge, intention and coercion. Each scenario had an average of 8.25 sentences and 71.25 words. The system generated an average of 87 facts for each scenario, 80 of which came from the composition and 7 of which were inferred by the discourse-level interpretation. 61 of those facts, or 70%, were used by the model of blame attribution in its reasoning process. The model inferred ordinal relations for the amount of moral responsibility between each of the 28 possible pairings of the eight agents in the four scenarios. Figure 13 shows these ordinal relations as a partially ordered graph. The agents are labeled either *chm* for chairman or *vp* for vice president followed by the scenario number 1-4. The number beneath each agent is the average blame attributed on a scale of 1-6 by the participants in Mao's survey. These numbers were also shown in Table 1, compared there with Mao's results.

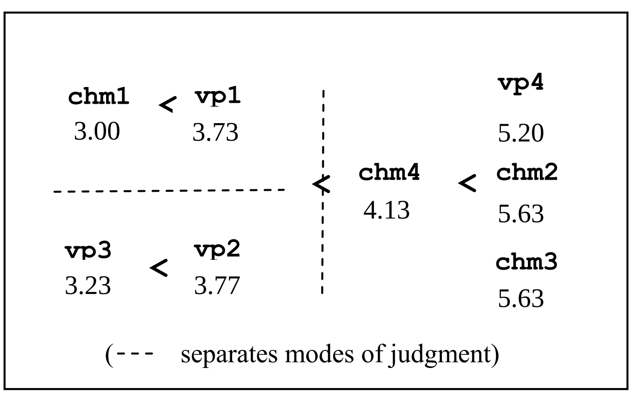

Figure 13: Ordinal constraints on responsibility

Using the representations generated by EA NLU, the model was able to correctly infer which agent in a pairing was attributed higher responsibility by the survey participants in 21 of the 28 possible comparisons. In 3 of the remaining comparisons, the model constrained the ordinal relation to depend on the relation between a pair of attribution variables, but no relation for those variables could be inferred from the text. In the remaining 4 comparisons, the model was insufficient to make a decision, due to gaps between the theories of Shaver and Weiner. This is a significant advance over the prior model.

3.2.4 Understanding cultural folktales

MoralDM has been evaluated against a study by (Dehghani, Gentner, Forbus, Ekhtiari, & Sachdeva, 2009) showing how cultural narratives can impact moral decision making. This study focuses on three well-known Iranian folktales that teach the moral value of sacrifice. The stories were selected through an internet-based pilot study using 199 Iranian subjects. Among other questions, subjects were asked to list the top ten cultural and religious moral stories they can think of. The three stories used were among the most referred to non-religious and non-political scenarios.

Three additional variants of each story were generated by the experimenters: surface changes (relative to the base scenario), structural changes and both surface and structural changes. The surface variants of these stories alter types (e.g. different instruments, different but related actions) while attempting to maintain the structure (e.g. causes, intentions) and ultimately the message of the story. The structural variants, in contrast, maintain the features of the actors, locations and props while altering the higher-order relations among the events that take place. Each base story concludes with a decision made by the protagonist; those decisions were not included in the variants. There were also variants reflecting change in protected values which do not add to the discussion here. Table 6 contains the original and QRG-CE versions of the base and variants for one of these stories. The other stories are shown in Appendix B.

	Original version	**QRG-CE translation**
Dehghan Fadakar		
Base	A farmer is returning home from a day of work carrying an oil lamp. He notices that as the result of a landslide, parts of a railroad just outside of a tunnel has been covered with stones. He walks passed the tunnel and realizes that a train is heading towards the tunnel. The farmer has two options, he can either try run to the station on time and inform the station manager and save his own life, or he can put his coat on fire, stand in the way of the train, risk his life and try to signal the train. He chooses the second option and saves the lives of many people.	A farmer is returning home from work and carrying a lamp. He notices that a tunnel has been blocked because of a landslide. He walks past the tunnel and realizes that a train is moving towards the tunnel. He has two options. The first option is, he can try to run to the station to inform the manager. This would save his own life. The second option is, he can stand in the way of the train and set fire to his coat to try to signal the train. This would risk his life. He saves many lives by choosing the second option.
Surface Δ	A man is going to work carrying a flashlight. He notices that as the result of an earthquake, a bridge has collapsed. He walks passed the bridge and realizes that a bus is heading towards the tunnel. He has two options: he can either try to run to the station on time, inform the station manager and save his own life, or he can use his flashlight, stand in the way of the of the bus, risk his life and try to signal the bus.	A man is going to work and carrying a flashlight. He notices that a bridge has collapsed because of an earthquake. He walks past the bridge and realizes that a bus is moving towards the bridge. He has two options. The first option is, he can try to run to the station to inform the manager. This would save his own life. The second option is, he can stand in the way of the bus and use the flashlight to try to signal the bus. This would risk his life.
Structural Δ	A farmer is returning home from a day of work carrying an oil lamp. He notices that as the result of a landslide, parts of a railroad just outside of a tunnel has been covered with stones. He walks passed the tunnel and realizes that a train is heading towards the tunnel. The farmer has two options, he can either try to run to the station on time and have the station	A farmer is returning home from work and carrying a lamp. He notices that a tunnel has been blocked because of a landslide. He walks past the tunnel and realizes that a train is moving towards the tunnel. He has two options. The first option is, he can try to run to the station to inform the manager. This would save his own life. The second option is, he can stand in the way of the

	manager reroute the train, or risk his life, by standing on the tracks, which will make him famous in his town and he would potentially receive a cash prize.	train and set fire to his coat to try to signal the train. This would cause him to be famous and receive money.
Surface and Structural Δ	A man is going to work carrying a flashlight. He notices that as the result of an earthquake, a bridge has collapsed. He walks passed the bridge and realizes that a bus is heading towards the tunnel. He has two options: he can either try to run to the station on time and have the station manager reroute the train, or he can use his flashlight, stand in the way of the of the bus, risk his life and try to signal the bus, which will make him famous in his town and he would potentially receive a cash prize.	A man is going to work and carrying a flashlight. He notices that a bridge has collapsed because of an earthquake. He walks past the bridge and realizes that a bus is moving towards the bridge. He has two options. The first option is, he can try to run to the station to inform the manager. This would save his own life. The second option is, he can stand in the way of the bus and use the flashlight to try to signal the bus. This would cause him to be famous and receive money.

Table 6: Iranian folktale variations from Dehghani et al. 2009

For each story, subjects in the study were given one of the variants at random and asked to decide what course of action the protagonist should take. The subjects were then questioned as to whether the variant reminded them of a story they already knew. Subjects in the Iranian population reliably retrieved the known base story for all the variants. In the base stories, the protagonist chooses personal sacrifice – a decision that is taught as being morally correct. This study demonstrated that subjects presented with a surface variant not only were reminded of the base story, but applied it to the variant scenario, choosing personal sacrifice as the protagonist in the base scenario did. They did so significantly more often than subjects presented with a structural variant. The latter group was still reliably reminded of the base story, but did not choose sacrifice. A control group with no prior knowledge of the base stories showed no such trend. This study provides evidence that cultural narratives inform moral decision making and that structural similarity is an important factor in the application to novel situations.

Understanding these folktales builds on the understanding of decision making scenarios discussed in section 3.2.2. A similar binary choice structure is presented in each variant, although here without the action/inaction distinction. It is also different in that both options seek the same primary outcome – saving people – but with different possible side-effects. The key addition deals with the intentions of the agent with respect to the choice being made. As in the attribution work described in section 3.2.3, intentionality is concerned with motivation for action combined with the willingness to bring about side-effects. In the example story, it is the farmer's intentional willingness to risk his life for the purpose of saving others that is considered a moral exemplar.

In the base version of the story, intention can be inferred from the choice that was made. The farmer was presented with a choice and chose one of the options, so the task model infers that he intended the previously mentioned consequences. It is notable that the side-effect, risking his life, was explicitly mentioned in the narrative prior to revealing his decision. By convention, even though the story does not say whether he knew of the danger, it can be justifiably assumed that he did because of the lack of evidence to the contrary. Domain reasoning could also conclude that the chosen action was dangerous, but the narrative provides the explicit clues in a way that eases the reasoning burden on the listener. The intentionality query in the base story results in the assertion:

```
(129)  (attributedValue Intention
            (ValueDuringFn (IntentionFn farmer57601 sel65687 save65200)
                           sel65687)
            Some)
```

where *farmer57601* is the protagonist, *sel65687* is the explicit act of choosing at the end of the story and *save65200* is the explicitly mentioned event of people being saved. It similarly returns the assertion:

```
(130)  (attributedValue Intention
            (ValueDuringFn  (IntentionFn farmer57601 sel65687 (DrsCaseFn DRS65686))
                            sel65687)
            Some)
```

where *DRS65686* is the case describing how he would risk his life, shown with its *possible-Historical* embedding in Figure 14.

```
Universe: farmer57601 option60183 …

(possible-Historical (DrsCaseFn DRS65686))

DRS65686:
   Universe: life70145 risk70113

   (isa risk70113 RiskTaking)
   (performedBy risk70113 option70183)
   (objectActedOn risk70113 life70145)

   (isa life70145 Living)
   (possessiveRelation farmer57601 life70145)
```

Figure 14: Partial DRS for "This would risk his life."

In the variations of the story, the protagonist has not yet made a choice at the end of the telling. Thus intentionality can only be inferred as implied for the hypothetical futures presented. If the character were to choose the first or second option, then he would be attributed appropriate intentions. This understanding is captured using the Cyc predicate *implies*, a binary

LogicalConnective that corresponds to the material implication operator of propositional calculus. The query:

(131) (implies (chosenItem ?sel ?option)
(attributedValue Intention
(ValueDuringFn (IntentionFn ?actor ?sel ?outcome) ?sel) Some))

attempts to prove that were a choice to be made, intention could be attributed based on the statement of known outcomes. This can be inferred from the recognition of the choice and consequences, here based on the representation of explicit statements such as shown in Figure 14.

All the variants for these stories exhibit a similar structure of choice and intention. This similarity is what prompts retrieval and alignment with the base story, modeled in MoralDM using analogical reasoning. In the surface variations, only types are altered – a bus vs. a train, a lamp vs. a flashlight. When the base story is retrieved as an analog, MoralDM concludes that the same choice, self-sacrifice, is appropriate in the new scenario. The difference in the structural change variant of the example story deals with the side-effect: from willingness to risk his life to the possibility of fame and fortune. The former representation is shown in Figure 14 while the latter is shown in Figure 15.

Universe: farmer18287 option20869 …

(possible-Historical (DrsCaseFn DRS-3448108765-26218))

DRS-3448108765-26218:

> Universe:
>
> (causes-ThingProp option20869 (DrsCaseFn DRS26216))
>
> DRS26216:
>
> > Universe: receive25955 money26010
> >
> > (renownLevel farmer18287 Famous)
> >
> > (isa receive25955 GivingSomething)
> > (givee receive25955 farmer18287)
> > (objectGiven receive25955 money26010)
> >
> > (isa money26010 Currency)

Figure 15: Partial DRS for "This would cause him to be famous and receive money."

MoralDM utilizes analogical reasoning to conclude that the high moral status of risking one's own life expressed in the base story transfers, by analogy, to being famous and receiving money. This conclusion is rejected by first-principles reasoning, invalidating the mapping. In essence, when given the structural variant the system is reminded of the base story but concludes that it is not applicable. Utilitarian reasoning then takes over and a non-sacrificial choice is made.

A similar reasoning process takes place in the other stories. All the base stories laud personal sacrifice for a morally valued outcome. These outcomes are identified, where they exist, by querying for the implication that it is true of a situation that it *hasHighMoralValue*. In the structural variant for the example story, *Dehghan Fadakar*, the moral value of the situation is altered by changing the actor's motivation. In the structural variant for the story *Pouryaie Vali*, the moral value of the situation is altered by changing the outcome from helping someone buy a house and get married to helping someone indulge themselves with expensive clothes. In the structural variant for the story *Hossein Sacrifice*, the moral value of the situation is altered by presenting a more effective non-sacrificial option. This is inferred by querying for a *moreProbableToAchieveOutcome* relation between two options of a choice. In all these cases, the compositional frame semantics of EA NLU are sufficient to cover the necessary semantic breadth to capture these distinctions. The additional queries used by MoralDM in this study are shown in Figure 16.

```
(queryForInterpretation 0 (intends ?agent ?situation))

(queryForInterpretation 0 (implies ?action (intends ?agent ?situation))

(queryForInterpretation 0 (implies ?antecendents (hasHighMoralValue ?situation))

(queryForInterpretation 0 (moreProbableToAchieveOutcome ?sel ?option1 ?option2))
```

Figure 16: Extension to MoralDM query model for Iranian folktales

To model this study, MoralDM was given a case library consisting of the base stories and presented with each of the variants. For two stories four variants were presented and for one story three variants were presented (leaving out the protected value change variant for which there was no human data). For each variant the model had to judge whether the protagonist should take the sacrificial choice or the choice that maximized personal utility. In these scenarios, the discourse-level representation generated by EA NLU had an average of 109 facts, only 3 of which were inferred. Of those 109 facts, only 38, or 35%, of them were used by MoralDM's first-principles reasoning module to make the decision matching human data. This low number is due to the fact that this experiment relied primarily on analogical reasoning. The 38 facts were used to infer the 3 additional facts, which were the facts directly queried for, and the entire set of 109 facts (on average) were used for analogical reasoning. Based on the representations generated by EA NLU, the system made choices for all eleven variants that matched the choices of subjects who were reminded of the base story.

3.3 Related work

3.3.1 Question answering and summarization

While early work in cognitive science involved a great deal of interest in stories and human memory (Schank & Ableson, 1977), systems that came out of that tradition were not used in cognitive modeling simulations. They were based on cognitive theories, but were not evaluated against psychological results regarding particular phenomena. Common evaluations instead included question answering (Cullingford, 1978; Wilensky, 1978) and summarization (Lehnert, 1981) tasks which have matured into specific challenges in the natural language community. However, these challenges have little to do anymore with cognitive modeling.

The Text REtrieval Conference (TREC) has involved a question answering track every year from 1999 through 2007. The types of questions have been expanded each year in scope and difficulty expanding from single fact answers to lists of facts and to a general question that requires extraction of all facts relevant to a target entity. The focus of the challenge from year to year has been on expanding the types of answers (e.g. from things to events), requiring increased integration of multiple document sources and expanding the scope of the corpus (Dang, Kelly, & Lin, 2007). The setup of the TREC QA challenge requires robust models that can process large corpora efficiently and extract bits of necessary knowledge. This encourages shallow processing approaches that are concerned with how lexical/syntactic features can be used as reasonable heuristics to identify answers. Successful approaches will often leverage semantic knowledge through semantic role labeling (Pradhan, Ward, Hacioglu, Martin, & Jurafsky, 2004) and semantic similarity of terms based on large-scale resources such as WordNet (Fellbaum, 1998). Recent approaches seeking to increase semantic depth to improve performance (Schlaefer et al., 2007) are just reaching this level of semantic understanding, a considerable distance from general cognitive modeling. Simlilar NLP techniques are used in the Text Analysis Conference (TAC) summarization challenge. Features such as named entities and parts of speech are combined with WordNet semantic distance metrics to score the importance of sentences within a larger text for distilling (Bawakid & Oussalah, 2008). This area has developed many practical information retrieval and text manipulation techniques. It has also served to better define what can and cannot be accomplished in a knowledge-poor environment, and shown that resources such as WordNet can fill an interesting gap in shallow semantics.

3.3.2 Discourse psychology

At the other end of the spectrum, work in discourse psychology has explored cognitive models of text understanding. This field of work is concerned with how a human reader constructs a referential situation model of what a text is about. In particular, research has focused on the generation of inferences during reading and the encoding of text in memory for retrieval. Despite the obvious connection between psychological models of the reading process and computational language understanding, little cognitive simulation has been done with these theories. As cognitive models of reading they are particularly interesting in relation to this work because they could both provide specific target models of reasoning and also inform more generally the contextual reasoning part of the semantic interpretation process.

Graesser, Singer and Trabasso presented a constructivist theory of inference during narrative comprehension (Graesser et al., 1994). This theory was specific to narrative understanding, being based on the hypothesis that narrative text has a close correspondence to everyday experience and that deeply embedded knowledge about actions, events, goals and emotional reactions play a key role in comprehension. Three critical assumptions, together the principle of *search after meaning* first proposed by (Bartlett, 1932), inform this theory: 1) that readers construct meanings that 1) address their goals, 2) are globally and locally coherent and 3) explain why actions, events and states are mentioned in the text. Based on these assumptions, this theory predicts that certain classes of inference are generated on-line as opposed to being generated during later retrieval and/or reasoning. These predictions were tested with numerous studies based on think-aloud protocols and reaction times (Singer, Halldorson, Lear, & Andrusiak, 1992; Suh & Trabasso, 1993). These predicted inferences could form the basis of an interesting cognitive model for reading, one that could be tested in EA NLU.

Experiments in (Kim, 1999) tested the hypothesis that the presence of *causal bridging inferences* in a story influences the perception of interestingness. Two versions of the same stories were constructed, one with five causally related sentences and one with the fourth sentence removed. Timing data supported the assumption that subjects presented with the second version would take longer to process the final sentence of the story than those presented with the first version, indicating the generation of causal bridging inferences. The subjects were also asked to rate the interestingness of the stories and the second versions were rated significantly more interesting than the first. Additional experiments provided evidence for discounting an alternative hypothesis that interestingness correlates with difficulty in comprehension, regardless of resolution. Kim's results provide another interesting constraint on the interpretation process that has not been computationally modeled.

Work by van den Broek (van den Broek, Lorch, Linderholm, & Gustafson, 2001) further explored the impact of reader goals on on-line inference. This work proposes *standards of coherence* as a framework for systematic understanding of the relationship between goals and inference. It suggests that there are a limited number of standards which, when selected, are maintained by the reader through inference. In the study, readers were given either a study goal or an entertainment goal. The main hypothesis was that these goals would affect the frequency of coherence-producing inferences. Readers with a study goal were expected to frequently make explanatory and predictive inferences based on stringent standards for coherence within the text. Readers with an entertainment goal, by contrast, were expected to more frequently make associative inferences to connect the text to their own experiences or values. The study provided evidence backing up this hypothesis. Readers with the study goal were stricter in their standard of coherence, taking more time and care to make explanatory connections between elements in the text, while the readers with the entertainment goal read quickly, made more associative inferences, and showed less concern for maintaining a coherent model. This is particularly interesting as a model of inference because it defines the implications of specific contextual factors.

The *construction-integration theory* (CI) (Kinsch, 1998) of reading has seen computational implementation (Mross & Roberts, 1992) and been evaluated as an explanatory theory for bridging inferences (Schmalhofer, McDaniel, & Keefe, 2002). CI posits two separate, repeating processing phases: construction and integration. The construction phase takes a new clause and instantiates or activates knowledge units as nodes in a multi-layer graph. The graph consists of a surface layer capturing the verbatim text and syntactic structure, a propositional layer for semantic meaning and a situational layer representing the referential state of affairs being addressed. The nodes are interconnected based on syntactic and semantic relations within and between levels. The integration phase simulates spreading activation among the network, with highly interconnected nodes becoming more highly activated and fringe nodes falling off. These activations have been shown to predict priming effects and what inferences will be made. Input to the CI program consists of hand-made parse trees, propositional logic and situation frames. EA NLU could be used to facilitate natural language input, and it is possible that this model would provide a profitable direction for considering how the salience of facts in a discourse changes over time.

3.4 Conclusion

The EA NLU system has been used to generate formal representations of stimuli for three separate cognitive simulations. In each case, the integration of knowledge-rich subcategorization frames, compositional frame semantics, choice sets, discourse representation structures and query-driven processing was sufficient to provide the range of semantic breadth required by the cognitive model. Several other simulations using EA NLU are in progress or used it as a guide to representation during early stages of development. The implementation of EA NLU as a working system has forced specific, consistent rules for semantic frames and semantic composition across these multiple models. This is an advance in reducing the tailorability of input to cognitive simulations that will be refined and increase in value as additional models are investigated. Further, the specification of these rules and their automatic enforcement reduces the knowledge engineering expertise required by experimenters. Subcategorization frames can almost always be added by copying the structure of existing frames, which is a much simpler task than representing complex nested compositions.

A related and critical measure of the usefulness of such a system is the ease of extension. This was discussed in general terms in the conclusion of chapter 2. In practice, I have seen improvements in effort through the process of working with these three sets of stimuli. It is difficult to measure this type of effort without an extensive user study, which would be premature for this system, so here I am limited to my own encoding experience. Adding subcategorization frames is a very straightforward process, given the high degree of common structure among frames. In those cases where a novel logical form in a semantic frame was required, it was typical that a novel semantic composition in the grammar was also required. The semantic compositions supported by the grammar are general-purpose across the sets of stories used in this work. Quantification, negation, implication, modality, utterance and clausal complements such as cause and purpose are very common semantic expressions. Extending these is not a trivial task, but neither is it a common task. Again based on a small sample, the number of term and frame additions for these scenarios is more than an order of magnitude greater than the number of complex semantic compositions. Moving forward with additional models, it would be useful to more formally study the types and frequency of extensions. It is certainly the goal of this work that the number of significant structural changes for each new set of stimuli trend down towards zero.

4 Limited Evidential Abduction for Disambiguation

In this chapter I describe my approach to disambiguation in facilitating natural language input to cognitive modeling and other knowledge-rich reasoning tasks. Disambiguation is one of the major challenges for computational natural language work. In my practical approach to language understanding, compositional frame semantics are deliberately separated from the general disambiguation task. Explicit ambiguities are maintained in the form of choice sets so that disambiguation can use pragmatic constraints and world knowledge. Cognitive models provide a well-scoped pragmatic context for language understanding, because the relevant aspects of personal, group and cultural context have already been examined and scoped by the psychological theories. In the same way, the impact of world knowledge is better understood.

I build on the insight, taken from several lines of research, that abduction is a natural mode of reasoning for understanding language. Language understanding deals with understanding how each subsequent utterance updates an incrementally constructed belief model. Those utterances carry both explicit and implicit knowledge. Grice framed this process in terms of *conversational implicatures* (Grice, 1975) that allow inference of the implicit knowledge communicated by each utterance. Making those inferences relies on making assumptions, which in turn support a reasonable explanation of the connection between the given and the new. That is, assuming unsaid content explains how each utterance makes sense. Abduction is a natural fit for this process because it provides a formal model of inference where a reasoning system makes assumptions in order to explain known and observed facts.

Once again, cognitive models provide a more specific, well-defined context. The pragmatics of interpretation is defined by the cognitive model, in the form of a query-driven task model. The ambiguities are defined by the explicit choice sets generated by the sentence-level compositional frame semantics. This chapter provides evidence that a limited form of abductive reasoning is an effective way to parsimoniously disambiguate those choice sets. I first discuss general abductive reasoning and the application of abduction to natural language understanding. Then my limited evidential abduction approach is described, including its implementation in EA NLU. I then show experimental results using one of the cognitive models described in chapter 3, and end with additional related work and concluding discussion.

4.1 Background

4.1.1 Abduction

Abductive reasoning is a nonmonotonic form of reasoning that searches for explanations for observed manifestations. It was first formalized by Peirce (Peirce, 1955) who concluded that if the presence of a “surprising fact” C could be explained by the presence of a fact A, this is reason to suspect that A is true. He also argues that the hypothesis that A is true is nothing more than a possibility, and that choosing the correct hypothesis out of a number of possibilities is “purely a question of economy”.

Abduction was first applied in Artificial Intelligence by Pople for a medical diagnostic task (Pople, 1973). Abduction is well suited to model-based diagnosis where observed manifestations (the problem) must be explained. This is probably the most common use of abductive reasoning in AI, although it has also seen use in, among other areas, plan recognition, planning, case-based reasoning, learning, vision and natural language understanding (Paul, 2000). Pople formulated abduction as linear resolution in a theorem-proving framework, stating that for a logical theory T and a formula ω which is the fact to be explained, a cause φ of ω in T is determined by three conditions:

(132) $T \wedge \varphi \models \omega$

(133) φ is consistent with T

(134) φ is abducible

In essence, abduction relies on a deductive relation between the fact to be explained and the explanation. Condition (134) allows for restriction in the set of abducible sentences (or predicates) so that dynamic restriction of the search space is possible. It is often chosen to cover all predicates in the language. This type of formulation is adopted in much research on abduction in AI, but it may be criticized as overly strict due to inheriting the constraints of deductive inference. Boutilier and Becher argue that a better model generalizes for defeasible entailment, variable plausibility of explanations (echoing Peirce's insight on possible hypotheses) and the necessity of belief revision (Boutilier & Becher, 1995).

Abduction has been shown to be much harder than classical inference, which fits the intuition that it takes the problem of deductive back-chaining and makes it far less constrained. In particular, it adds a significant second source of complexity in deciding which hypotheses are preferred. Bylander et al. analyzed a very general definition of abduction that covers logical formulations as well as set-covering (Reggia, Nau, & Wang, 1983) and baysian belief revision (Pearl, 1987). They showed that the general case is NP-hard because of the nature of the problems, regardless of the representation or algorithm being used (Bylander, Allemang, Tanner, & Josephson, 1991). Restricted classes have been found that are tractable in polynomial time, but involve assumptions unlikely to hold true in real domains (Bylander et al., 1991; Eshghi, 1992; Selman & Levesque, 1990).

4.1.2 Abduction and Natural Language Understanding

Early work in natural language understanding produced the insight that interpretation is a process of forming a reasonable explanation of the meaning of the utterance, in context, based on prior knowledge and assumptions. Several lines of research have followed this insight using the framework of abductive reasoning. This has provided more uniform, well-understood representations of knowledge and inferential processes that have been more easily compared in the community. I will discuss the particular formulation of (Hobbs et al., 1990) here, and expand on other approaches in the related work section 4.5.

In the application of abduction to natural language understanding, the manifestation is the surface form of an utterance (in context) and the explanation is what the speaker intended to communicate (in the hearer's judgment). Taking an example from (Hobbs et al., 1990), consider the sentence:

(135) The plane taxied to the terminal.

This approach creates an underspecified logical form of the surface syntax (essentially semantic role labeling) which is:

(136) $(\exists x,y) plane(x) \wedge taxi(x,y) \wedge terminal(y)$

The predicates *plane*, *taxi* and *terminal* are *lexical predicates* which represent the descriptive usage rather than a type in the world. Given the following set of axioms in a knowledge base:

(137) $(\forall x) airplane(x) \supset plane(x)$

(138) $(\forall x) wood\text{-}smoother(x) \supset plane(x)$

(139) $(\forall x,y) move\text{-}on\text{-}ground(x,y) \wedge airplane(x) \supset taxi(x,y)$

(140) $(\forall x,y) ride\text{-}in\text{-}cab(x,y) \wedge person(x) \supset taxi(x,y)$

(141) $(\forall x) airport\text{-}terminal(x) \supset terminal(x)$

(142) $(\forall x) computer\text{-}terminal(x) \supset terminal(x)$

(143) $(\forall z) airport(z) \supset (\exists x,y) airplane(x) \wedge airport\text{-}terminal(y)$

this approach proves the surface form (136) by assuming the existence of an *airport* (which entails the existence of an *airplane* and an *airport-terminal* by axiom (143)) and a *move-on-ground* event involving that *airplane* (which entails a *taxi* event by axiom (139)). That is, the sentence (135) is explained by assuming that the speaker is describing an airplane at an airport moving towards an airport terminal. An alternative explanation assumes the existence of a *wood-smoother*, a *ride-in-cab* event, a *person* doing the riding and a *computer-terminal*. The key distinction between several approaches is how they evaluate the quality of one explanation over another.

Hobbs' formulation of interpretation by abduction demonstrates that it can extend to syntactic and discourse constraints. Rather than prove an underspecified logical form such as (136), the system can prove that the surface utterance (135) is a sentence. Grammar rules can be encoded as axioms that prove, for example, that the existence of a noun phrase (NP) followed by a verb phrase (VP) entails a sentence. Those axioms can be augmented with semantic information such that the sentence is entailed by the syntactic constituents together with their semantic contributions (the elements of the logical form (136)). In a similar manner, Hobbs shows that this approach can be extended from proving sentences to proving the discourse structure of a sequence of sentences. He proposes a discourse structure based on binary *coherence relations* between utterances, but argues that any well-formed theory (Hobbs, 1985; Hovy, 1988; Mann & Johnson, 1986) could apply. Hobbs argues that this approach is particularly powerful because it integrates syntactic, semantic and pragmatic constraints, not giving one type of assumption prominence over another. This approach uses Stickel's *weighted abduction* (Stickel, 1989) scheme to determine the best explanation for the set of possible hypothesis. Every expression to be proven is given a cost for assuming it, and every axiom in the knowledge base is annotated with weights indicating the cost of assuming each antecedent as a function of the cost of the consequent. Thus each explanation is measured by the sum of the costs of the assumptions that entail it, and the explanation with minimal cost is selected. Further, redundant expressions in the explanation are factored together and assigned the lesser cost. This ensures that explanations featuring assumptions that entail multiple goal expressions, such as the airplane taxiing at the airport, are preferred.

4.2 Limited evidential abduction

In this work, I have used the task models presented by cognitive models as a pragmatic guide to understanding. The queries presented by each model capture the pragmatic language understanding task being performed. In abductive terms, the manifestation is a particular interpretation that is being sought, and the explanation is whether that interpretation can be supported. This is not to claim that people in general, or even the subjects in the source psychological studies, approach a particular text with a particular agenda. Rather I argue that a model of language understanding cannot be said to ascertain that a particular interpretation is correct without pragmatic context. My investigation here is limited to the assumption of a particular context in any given case: that provided by the cognitive model being explored.

I use abduction to model interpretation as the search for an explanation that justifies a particular interpretation, one defined by the pragmatic context of a cognitive model. However, there are significant issues with abduction as a general approach. From a practical standpoint, it is generally intractable. This problem is exacerbated because broad world knowledge is necessary, but previous approaches have stopped short of specifying how the process interfaces with large-scale knowledge without combinatorial explosion. General abduction presents another problem in that it is a global proof. Certainly one cannot imagine reading an entire paragraph, much less a chapter or book, before reasoning about what the first sentence meant. Further, there is a significant and very open question regarding evaluation of the quality of explanations. Global cost metrics, whether hand represented or based on Baysian probability theory, have been criticized as inflexible because they do not use the particular context of the interpretation (Norvig & Wilensky, 1990). This is exacerbated by the fact that interesting narratives are often pointedly about uncommon or improbable occurrences. Coherence within the explanation, such as measured by factoring redundancy or explicitly valuing connectedness, is a good metric for the plausibility of an explanation, but suffers the same problem.

I propose *limited evidential abduction* to address these issues. Prior approaches have allowed any fact to be assumed, creating a large space of possible explanations and placing the burden on evaluation to control the search. By contrast, this approach starts with an empty abducible set and selectively adds to it. Specifically, those facts for which some form of evidence outside the proof can be found are considered reasonable assumptions and added to the set. This is not mutually exclusive with global or proof-internal heuristics, but at this point they have not been needed. For the problem of disambiguating among the choice sets generated by EA NLU's compositional frame semantics, the choice sets themselves are considered evidence for assumption. Each choice set is a mutually exclusive set of reasonable assumptions based on the sentence that generated it. Thus any choice that does not conflict with previously established choices can be assumed. No additional heuristics for choice preference are needed, only the domain theories already necessary for understanding a scenario with respect to a particular cognitive model.

4.3 Abduction in EA NLU

4.3.1 Abductive query

I have implemented limited abduction in the FIRE reasoning engine for use with EA NLU. Abduction is invoked by calling the function *abductive-query* which is passed a pattern in CycL and a *context* in which to query. This context corresponds to a case in FIRE's LTMS-based working memory. Importantly, axioms are contextual in the knowledge base, being stored in microtheories as other facts are. This means that only those axioms that are true in the specified context, based on microtheory inheritance, are used in the proof.

An *abductive-query* proceeds as a normal back-chaining query in FIRE, first attempting to retrieve ground facts unifying with the query expression from working memory or the knowledge base, then retrieving applicable axioms and recursively querying for the antecedents. Each query returns zero or more sets of valid, consistent variable bindings for the query expression in the given context. Those sets are the list of sets returned by retrieving ground facts concatenated to the list of sets returned by querying all applicable axioms. Back-chaining axioms in FIRE are expressed as PROLOG-style rules with order-dependent antecedents. When querying a set of antecedents a cumulative set of variable bindings is maintained. Prior to querying a particular antecedent, those bindings are applied to the antecedent expression. If that query returns no sets of variable bindings, that branch of the query fails. If it returns one set, the cumulative set of bindings is updated as the conjunction of the current cumulative set and the new set. If it returns multiple sets, the query branches by the number of sets, and for each branch the cumulative set is updated with the new set for that branch. For all successful branches, the antecedent facts with the cumulative bindings applied are asserted in the LTMS to imply the consequent of the axiom with those bindings applied. The list of those successful cumulative binding sets is returned for the query of that axiom.

Limited abduction is implemented by adding axioms to the environment with antecedents of the form:

(144) (abductiveAssumption <expression>)

When an expression of this form is queried, it invokes a call to *make-abductive-assumption*, a method which selects on the predicate of *expression*. This method is responsible for determining whether there is evidence for assuming *expression* and whether it is consistent with the current logical environment. Specializing on the predicate is useful for EA NLU choice sets, but clearly other criteria could apply. If the system judges that it is reasonable to assume *expression*, which must be ground at the time of the query, it is asserted in the LTMS that:

(145) (implies <expression> (abductiveAssumption <expression>))

and the query for *(abductiveAssumption <expression>)* successfully returns. This creates a situation where if *expression* were to be assumed true in the LTMS, it would also make *(abductiveAssumption <expression>)* true. However, *expression* is not assumed true at this point because assumptions in different branches of the proof may conflict with one another. Instead, it is returned as a part of the variable binding set using the special variable *?abduction-asms*. When a query returns a binding-set with *?abduction-asms* bound, it indicates that the rest of the set satisfies the query only if the expression bound to *?abduction-asms* is assumed to be true. This expression is stored in disjunctive normal form (DNF) and may represent more than one set of facts that entail the result of the query.

When a set of antecedents is being queried, in order, a cumulative *?abduction-asms* must be maintained. Consider an axiom in the PROLOG-style form:

(146) (<== (q ?x)
(p1 ?x)
(p2 ?x)
…)

where *(p1 ?x)*, *(p2 ?x)* and some additional antecedents imply *(q ?x)*. Because the antecedents are ordered, *(p2 ?x)* will only be queried if *(p1 ?x)* returns one or more bindings for *?x*.

However, with an abductive query, *(p1 ?x)* may return a binding for *?x* combined with a binding for *?abduction-asms*. That DNF expression is passed into the query for *(p2 ?x)* as the current *abduction environment*. If the query for *(p2 ?x)* also leads to a possible abductive assumption, it will only be considered if it does not conflict with the abduction environment. Conflicts are detected when the conjunction of the potential assumption and the abduction environment resolves to *nil*. The resolution process is standard, with calls to the method *conflicting-asms* which is responsible for enforcing exclusivity constraints between pairs of expressions. If the query for *(p2 ?x)* returns a binding for *?abduction-asms*, that expression is added to the abduction environment for the subsequent antecedent queries. That merged environment is also returned as a binding for *?abductive-asms* for the query to *(q ?x)*, if that query succeeds.

In EA NLU, the predicate *selectedChoice* is used to relate a particular selection to a choice set. To support limited evidential abduction, EA NLU specializes the *make-abductive-assumption* and *conflicting-asms* methods on this predicate to enforce the following rules. First, each choice set is a mutually exclusive set. Second, all non-parse tree choices are entailed as choices by the selection of one or more parse trees. Therefore if a parse tree has been selected, only choices that it entails may be selected. Conversely, if any choice has been selected, only parse trees that entail it as a choice may be selected.

When an abductive query completes, it returns a set of variable bindings that can be applied to the query form to produce the ground expressions that can be proved. For any such expression where there is no binding for *?abduction-asms,* the expression is already entailed in the LTMS. If the query returns one possible true expression together with a binding for *?abductive-asms*, then that set of assumptions can be assumed true, entailing the expression and providing an explanation for it. This may involve assuming a disjunction. The LTMS makes it possible to ascertain what facts in the explanation beneath the proven expression are definitely true and what sub-trees are only known to be part of a true disjunction. The latter case corresponds to the competing hypotheses found in general abduction. Limited abduction allows for a conclusion to be reached regarding the query form without necessarily disambiguating those disjunctions, but also identifies them (together with the justification structure in the LTMS that is used in other evaluation metrics) for evaluation if the task calls for it. If the query returns multiple possible true expressions, the abductive query has proven multiple conclusions with multiple explanations. Depending on the pragmatic context, the query may require a single answer or not. In the former case, a measure of the quality of the explanation may be applied, such as minimizing the number of assumptions, maximizing connectivity, minimizing cost probabilities or maximizing a proof-external scoring function. In the latter case, the set of answers may be divided into consistent subsets (where no assumptions in the set conflict) and then quality measures applied. In the work described in this chapter, I utilize a simple minimization of assumptions metric. A more advanced scoring function is discussed in the next chapter.

4.3.2 Discourse interpretation as abduction

The final step in EA NLU's compositional frame semantics process is transforming the semantic translation for each parse tree from predicate calculus into nested DRS, as discussed in section 2.3.5. However, those translations contain embedded choice sets as explicit points of ambiguity. Thus the output of the process is not a single DRS, but a set of axioms which entail the facts that make up the possible DRS structures based on what choices are selected. This set of axioms together with the explicit choice sets form an environment where limited evidential abduction can be applied to prove a pragmatic interpretation.

In order to create these axioms, EA NLU does a depth-first walk through the predicate calculus tree for each parse tree, branching at each choice set. An antecedent stack is maintained, beginning with assertions of the sentence position in the discourse and the selection of the particular parse tree. Whenever a branching point is reached, the selection of the choice for that branch is pushed onto the antecedent stack. Whenever a quantifier or logical connective is reached, a new DRS is asserted as implied by the conjunction of the expressions on the antecedent stack. In the case of a quantifier, the quantified variables are implied to be in the universe of that DRS, as described in section 2.3.5. Whenever a first-order fact is reached, it is asserted to be implied by the conjunction of the expressions on the antecedent stack as well. This process is combinatorial in the number of possible nesting structures, but importantly not in the number of total choices. The majority of the choice points come from multiple semantic frames, and the critical scaling question for the system is how it deals with the unlimited number of semantic frames that could be added. The vast majority of existing frames do not involve nested clauses, and these choices are independent of each other with regard to the complexity of the DRS transformation.

Discourse interpretation using abduction proceeds as follows. For each sentence in a discourse, compositional frame semantics are used to build the set of axioms and choice sets described above. The queries for the chosen task model are then queried as described in section 2.4, but using *abductive-query*. In most tasks, the appropriate model for abductive proof is one or more conjunctions of query expressions. This attempts to prove that the entire conjunction can be true for some set of assumptions. A particular strategy must be employed for choosing between competing answers if that is appropriate for the task. Finally, the assumptions for the chosen answer(s) are made, disambiguating some number of the existing choice sets. These entail some number of sentence-level facts which are merged with the discourse-level DRS as described in section 2.4.3. As the abductive queries may have selected choices from any prior sentence in the discourse, the merge process is executed for each updated sentence. The DRS representation makes such incremental updates no more complex than the initial merge for a sentence.

An abductive proof spanning the discourse could simply be run once after all the sentences have been processed. However, in the case of multiple queries, it is possible that some can be proven by less than the full set of sentences. In that case, earlier sentences are partially disambiguated, potentially reducing the complexity of the later queries.

4.3.3 An example of limited evidential abduction

Consider the Ritov and Baron dam scenario shown in Figure 17 (Ritov & Baron, 1999). This is one of the scenarios used in the MoralDM experiments described in chapter 3.

> Because of a dam on a river, 20 species of fish will be extinct. You can save them by opening the dam. The opening would cause 2 species of fish to be extinct.

Figure 17: QRG-CE version of the Ritov and Baron dam scenario

The compositional frame semantics generates parse tree, frame semantics, scoping and reference choices sets for each sentence in the discourse. As an example, the second sentence generates a single parse tree and three ambiguous frame semantics choice sets. These are shown in Figure 18, Figure 19 and Figure 20.

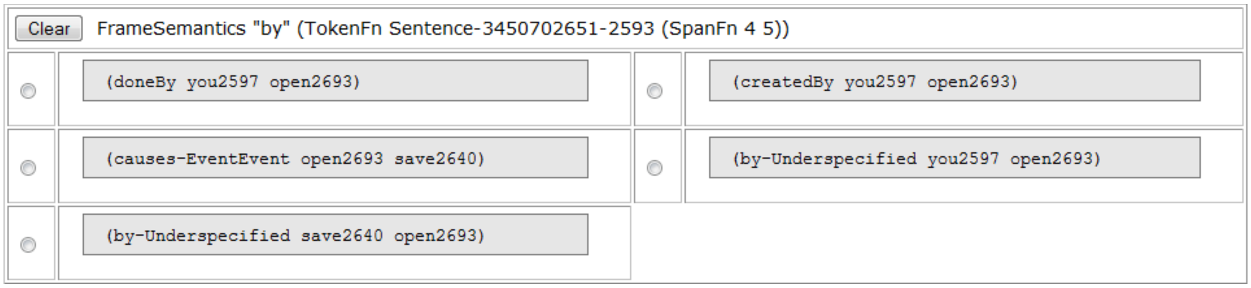

Figure 18: Semantic frame choices for "by"

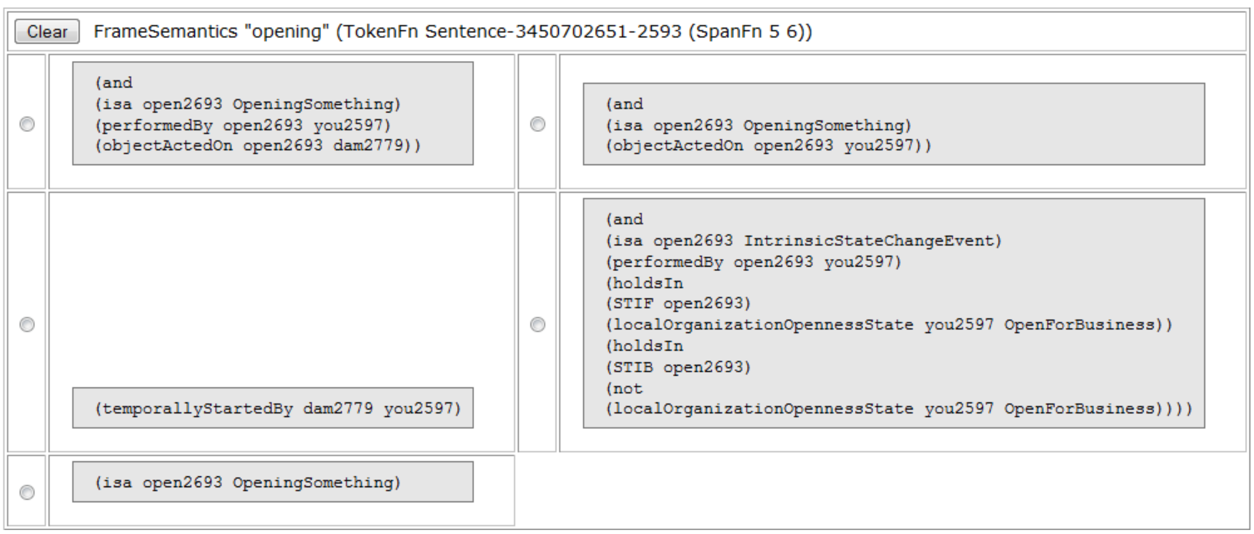

Figure 19: Semantic frame choices for "opening"

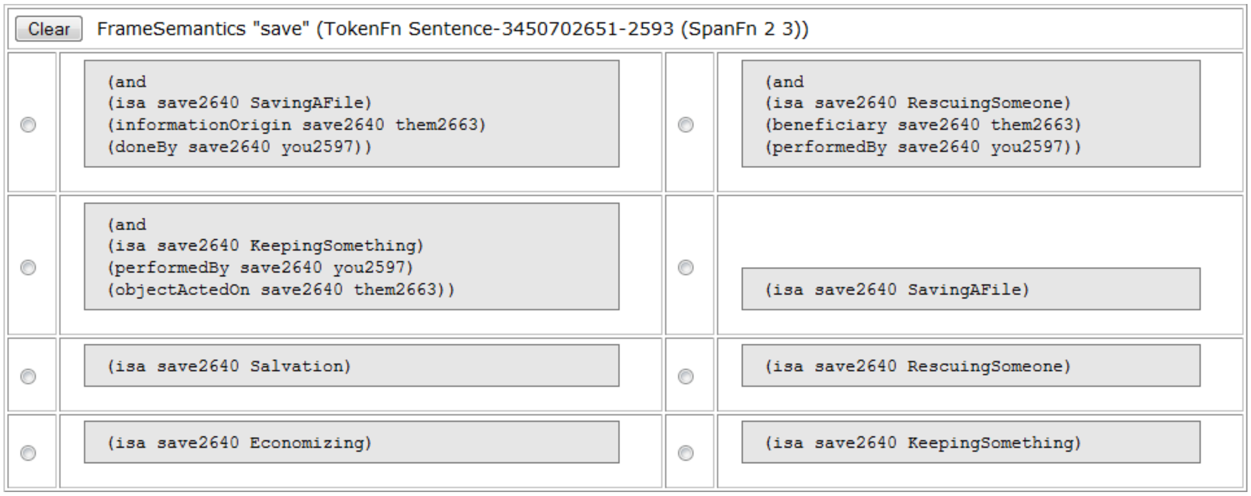

Figure 20: Semantic frame choices for "save"

In the experiment described in this chapter, scoping collisions between modals and quantifiers resulted in a point of ambiguity. Thus this sentence also generated a scoping choice set. The phrase "save them by opening the dam" translates to the (abbreviated) quantified logical form:

(147) (thereExists (TheList them2663 save2640 dam2779 open2693)
(and (isa save2640 RescuingSomeone)
(beneficiary save2640 them2663)
…)))

When the semantic translation (147) is composed with the auxiliary modal "can" to form the constituent phrase "can save them by opening the dam", there is a scoping ambiguity between the modal operator *can* and the existence of the four discourse variables *them2663, save2640*, *dam2779* and *open2693*. The entities referred to by these variables may exist only in the hypothetical future indicated by *can* or in the scope outside the modal. Some heuristic simplifications were applied in this version of the system. The pronominal reference *them2663* and the definite reference *dam2779* are scoped at the highest level in the sentence level DRS, given their referential nature. With the additional simplification that the existential variables are considered together, the choice set reduces to two choices:

(148) (outscopes possible (thereExists (TheList save2640 open2693)))

(149) (outscopes (thereExists (TheList save2640 open2693)) possible)

These simplifications were useful for the scope of this particular experiment, but were discarded with the change to scoping described in section 2.3.4.8. The compositional semantics now assume only the weakest scoping assertion (in this example that the two events exist at least in the hypothetical future), leaving further distinctions to the contextual query-driven reasoning.

The second sentence also generates three open choice sets for the three references "you", "them" and "the dam". Table 7 gives the number of choice sets and number of choices for each type of choice set for all the sentences in the example story in Figure 17. Since each choice is entailed by a particular parse tree, the choice sets are presented per parse tree to give a better indication of the total choice space.

Sentence/Parse Tree	**Frame Semantics Choice Sets**	**Frame Semantics Choices**	**Scoping Choice Sets**	**Scoping Choices**	**References**
sentence1, parse tree 1	4	11	2	4	0
sentence1, parse tree 2	4	8	4	12	0
sentence2, parse tree 1	3	15	1	2	1
sentence3, parse tree 1	5	15	1	2	1
sentence3, parse tree 2	5	16	1	2	1

Table 7: Choice sets and choices for the Ritov and Baron dam scenario

As described in chapter 3, the MoralDM task model queries to identify violations of sacred values and to identify choice and consequences. These queries are expressed here as conjunctive queries:

(150) (violationOfSacredValue ?outcome)

(151) (and (isa ?sel SelectingSomething)
(choices ?sel ?action)
(choices ?sel ?inaction)
(different ?action ?inaction)
(isa ?inaction Inaction)
(causes-PropSit (chosenItem ?sel ?action) ?outcome)
(causes-PropSit (chosenItem ?sel ?inaction) ?outcome2))

In this example, violation of a specific sacred value is inferred for each instance of extinction of some number of biological species. Those events fill the *?outcome* role in the query for the first and third sentences, respectively. The semantic frame choices for the term "extinct" are shown in Figure 21.

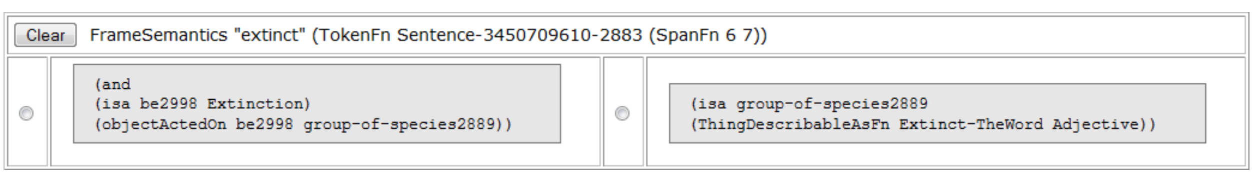

Figure 21: Semantic frame choices for the "extinct"

As opposed to in the earlier experiments, described in chapter 3, this choice has not been made via user intervention for the back-chaining query to discover. Rather, the choice involving an *Extinction* event must be abductively assumed (along with other assumptions about *group-of-species2889*) in order to prove that this scenario involves the violation of a sacred value.

4.4 Evaluation

The evaluation of MoralDM described in (Dehghani et al. 2008) demonstrated the capability of EA NLU to meet the understanding requirements of the reasoning model. Each scenario from the source experiments was rendered in QRG-CE and input to EA NLU. The system generated explicit ambiguities which were presented to the experimenter for manual disambiguation. Given this intervention, the system was able to produce logical representations sufficient for MoralDM to model human decision-making outcomes.

Here we present an evaluation of limited evidential abduction for automatic disambiguation within these established constraints. The four scenarios from Ritov and Baron used in the prior experiment (given in chapter 3, Table 2) are used. Each scenario is taken in its simplified form and processed by EA NLU. The system then queries for the same set of facts that MoralDM queries for use in its first-principles reasoning module, as described in section 4.3.3. The query is handled as an abductive proof which disambiguates the choice sets from the compositional semantics. Table 8 gives the number of ambiguous choice sets (parse trees, frame semantics, quantifier scope and references) in each scenario. For the three closed sets, the average number of choices is given in parenthesis.

	Parse Trees	**Frame Semantics**	**Scoping**	**Reference**
Scenario1 (dam)	3 (1.7)	13 (4.5)	5 (2)	4
Scenario2 (convoy)	2 (1)	15 (4)	5 (2)	5
Scenario3 (finance)	3 (1)	6 (5.2)	3 (3)	4
Scenario4 (logging)	3 (1)	13 (3.5)	3 (2)	5

Table 8: Choice sets for Ritov and Baron scenarios (average number of choices per set given in parenthesis)

Table 9 contains the same figures for unresolved choice sets after EA NLU performs the abductive proof of the MoralDM query-based task model. These are the choice sets that the system did not need to resolve in order to prove the necessary facts – they are considered spurious in the context of this task. Because there are constraints between choices, often the system reduced the available choices even when the set itself containing them was not resolved.

	Parse Trees	**Frame Semantics**	**Scoping**	**Reference**
Scenario1 (dam)	0	5 (2.8)	1 (2)	1
Scenario2 (convoy)	0	6 (3.5)	2 (2)	2
Scenario3 (finance)	0	1 (2)	0	1
Scenario4 (logging)	0	2 (3.5)	0	2

Table 9: Unresolved choice sets for Ritov and Baron scenarios (average number of choices per set given in parenthesis)

Table 10 presents the complexity space for each scenario. The worst-case number of random choices to satisfy the query is compared with the number of assumptions made by the abductive proof. The latter includes every time in the proof that the system checks to see if a fact can be or is already assumed. The space of unresolved choices is also provided. That table is sorted according to increasing number of worst-case random choices.

	Total Choices	**Abductive Assumptions**	**Unresolved Space**
Scenario3 (finance)	$6.27\text{x}10^4$	$6.21\text{x}10^3$	2
Scenario4 (logging)	$7.17\text{x}10^6$	$1.05\text{x}10^5$	10
Scenario1 (dam)	$7.63\text{x}10^7$	$1.01\text{x}10^4$	288
Scenario2 (convoy)	$6.94\text{x}10^8$	$8.33\text{x}10^4$	768

Table 10: Choice space vs. abductive assumptions

Figure 22 shows a logarithmic graph of the total choices vs. the abductive assumptions and the unresolved space, sorted by increasing number of total choices.

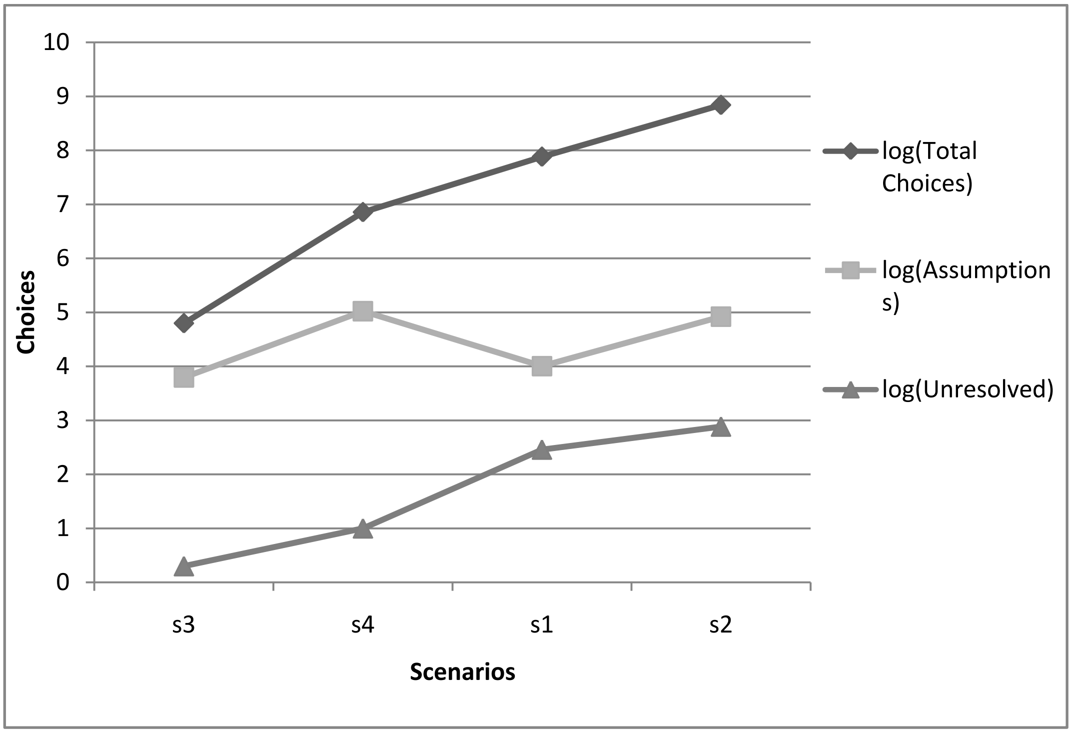

Figure 22: Choice space vs. abductive assumptions (logarithmic)

In all four scenarios the abductive proof is able to provide the facts requested. The number of choice actions taken by the solver is between one to four orders of magnitude smaller than the space of possible choices. What is most notable though is that as the space increases, the number of unresolved (task-irrelevant) choices increases while the number of assumptions does not. This demonstrates that this approach is able to make the necessary choices without suffering as the number of unnecessary choices increases. In this context this is particularly important as each additional sentence increases complexity regardless of whether the added ambiguities are task-relevant.

The types of ambiguities that the system did not resolve were largely surface distinctions in entity types. There were, for example, several ways of representing “species of fish” that did not impact this decision making task. Almost all scoping ambiguities were resolved by the system. Since hypothetical futures were a central part of understanding the decision, this is not unexpected.

As noted earlier, the scoping ambiguity choices were simplified in EA NLU following this experiment. Figure 23 shows the data with the impact of the scoping ambiguities removed from the worst-case random choices and unresolved choices. The same encouraging trends can still be seen in this case.

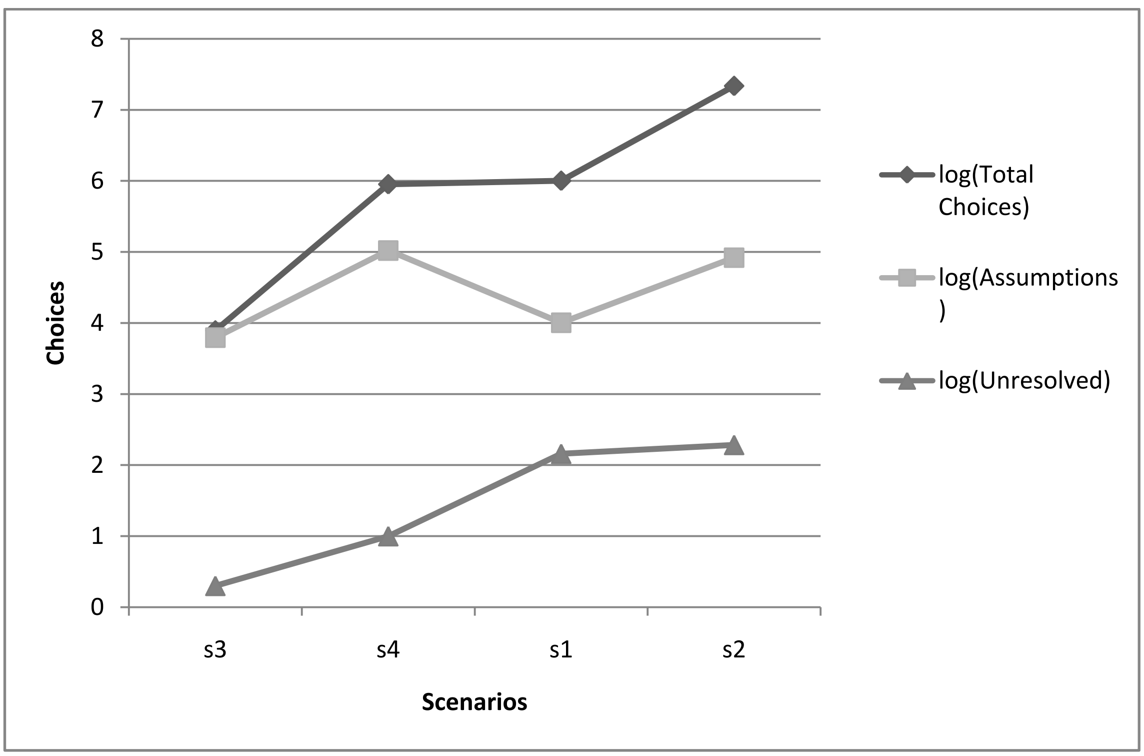

Figure 23: Choice space without scoping ambiguities

4.5 Related work

Cox and Pietrzykowski present a formulation of abduction in a theorem-proving framework (Cox & Pietrzykowski, 1986) that defines the abducible set based on notions of causality. In this model, only *basic* hypotheses may be abductively assumed, those that act as a cause rather than explaining an effect in terms of other causes. Mechanically, these are so-called “dead-end” branches of the proof that do not allow for a more specific hypothesis to be reached. This approach provides an interesting limit case by requiring maximal specificity, but clearly the most specific explanation is not equivalent to the best explanation. The authors, and others, point out that this approach both explores and relies on the incidental organization of the knowledge base.

Wilensky proposed a class of inferences referred to as *concretion* for understanding natural language (Wilensky, 1983). Concretion involves taking a certain relation or feature, in particular the underspecified relations and features commonly explicit in natural language, and inferring a more specific relation or feature. Wilensky suggests that the *primal* content of an utterance is an abstraction that covers all possible interpretations (Wilensky et al., 2000) and points out that according to (Searle, 1979) such an abstraction is never the actual meaning intended. He gives the example of the word “on” in “The cat is on the mat” having a primal interpretation of some manner of support. This is similar to the use of underspecified forms in logic, such as the set of underspecified predicates in Cyc (e.g. *on-Underspecified*) that are not intended for reasoning. Concretion in this example is inferring a more specific relation that is, in Cyc terms, a *genlPreds* of the primal relation. While Wilensky’s *UNIX Consultant* project (Wilensky et al., 2000) used a special-purpose concretion engine, Norvig proposed a general-purpose inferential system based on marker passing that included concretion (divided into type and relation concretion) as a class of inference (Norvig, 1987). Later work moved away from marker passing to general logic programming as a framework.

Charniak and Goldman (Charniak & Goldman, 1988) proposed a logic for semantic interpretation implemented in the Wimp2 system via forward chaining in an assumption-based truth maintenance system (ATMS) (deKleer, 1986). This system relied on frame knowledge, expressed as facts and axioms, to make explanatory assumptions by invoking frames and seeking to fill roles. This approach can be seen as logically equivalent to abductive back-chaining, with the required knowledge and axioms taking an alternative form.

Later work by Charniak and Goldman used a Baysian belief network (Charniak & Goldman, 1989) to represent the influences between events and objects grounded in the real world. These probabilities are used to address the problem of evaluating the best explanation for the manifest text. However, while they may attain a more principled grounding than the per-rule weights of Hobbs et al, these probabilities are still determined globally, independent of context or the story being heard. Ng and Mooney (Ng & Mooney, 1990) argue that coherence within the proof is the most significant metric of explanation quality in abduction. Their approach does not rely on a priori weights or probabilities but rather measures each explanation in terms of the connectedness of the network between assumptions and manifestations. Norvig and Wilensky criticize both of these approaches, as well as Hobbs et al, as oversimplifying the problem of evaluation while remaining computationally problematic (Norvig & Wilensky, 1990). They argue that a combination of a priori probabilities and textual coherence is necessary, as well as a better heuristic guide to make the problem tractable.

4.6 Conclusion

This chapter described my approach to disambiguation in the practical EA NLU framework. This approach provides a clear definition of the pragmatics of understanding by relying on cognitive models. It also provides a clear disambiguation task in terms of the explicit choice sets generated by the compositional frame semantics.

This approach uses abductive reasoning to formulate the problem, treating the desired pragmatic interpretation as the manifestation and the choice sets as the abducible set. This represents a concrete task for abduction as language interpretation, of which there have been very few, and introduces practical constraints to make the problem tractable. By limiting abduction to the choice sets, the pressure for breadth is placed on the semantic frames. This is a very practical approach because the complexity of the abductive proof depends on the query model and the axiomatization of the domain theories rather than on the number of choices. As in the case of compositional frame semantics, the biggest scaling challenge is in the number of semantic frames that can be added to the system. I have provided evidence here that this approach is able to scale in this dimension.

This approach does not address the complexity of the domain axioms needed for the abductive proof. These axioms are required if the system should have any capability of inferring the abstractions of the model (e.g. choice, intent) from the explicitly stated events of a narrative scenario. Engineering such domain theories is difficult and complex work, but it is required for deep reasoning about such models. This approach to disambiguation does not add to that requirement, but leverages what must already be there.

The global nature of abductive proof in natural language remains a significant problem. It is not just scaling in the number of utterances that is problematic, but also in the complexity of the proof. As the interpretation of a cognitive model becomes more and more complex, involving more and more entities and events, this approach will also suffer from scaling problems. This points to the need for a heuristic guide that operates incrementally as subsequent utterances are observed. In the next chapter I will present such an approach.

5.0 Narrative Functions as a Heuristic for Relevance

This work uses cognitive models to provide a pragmatic context for knowledge-rich language understanding. These models each define a context in which particular facets of a narrative are considered relevant to the reader, and the task of inferring those facets guides and constrains the semantic interpretation process. This chapter addresses a broader task, the use of narrative to illustrate a particular point the fables. This task is an instance of telling a story in order to communicate a point, which is a fundamental aspect of narration. This provides evidence that the theory and practical implementation created for this task have promise towards more general narrative understanding. There are two motivations. First, testing EA NLU and my practical approach to language understanding on increasingly broader reasoning tasks provides evidence for generality. Second, as described in chapter 4, global abductive proofs have scaling issues, pointing to the need for incremental, heuristic proofs that work on a sentence-by-sentence basis. This chapter describes such a set of heuristics, and demonstrates their application to identifying the meaning of a narrative.

This chapter contributes a theory of *narrative functions*, which are communicative acts performed by the narrator in the process of narration, and a task model that uses them for semantic interpretation. I claim that the task of inferring these functions in a narrative can serve as an incremental heuristic for inferring what is relevant about the story. The task is formulated as a set of expectations that the functions will appear in the narrative. It is implemented in the EA NLU discourse-level interpretation process as a set of query forms to be abductively proven for each sentence added to the discourse. The sufficiency of this reasoning task to guide interpretation towards intended meanings of a narrative is evaluated using a set of Aesop's fables.

I start with a discussion of prior research on general narrative understanding which my approach draws upon. I then present my theory of narrative functions and the task model of expecting them, showing how it is implemented in EA NLU. This is followed by application to a set of Aesop's fables, including detailed analysis of how expecting narrative functions guides interpretation of those fables, how the morals of the fables capture the meaning they are intended to communicate, and an evaluation of the system on the reasoning task of identifying applicable morals. Finally, I contrast related approaches and conclude with a general discussion.

5.1 Background: understanding language and narrative

5.1.1 Linguistic communication

Human linguistic communication involves a significant amount of inference on the part of the hearer. Where the speaker intends to communicate a certain set of conclusions, effective communication requires that the hearer draws those very conclusions. It is an imprecise process, yet astoundingly flexible and effective. Most current research explores this pragmatic phenomenon within the co-operative framework set out by Grice (Grice, 1975). Grice argued that the hearer is not merely decoding the content encoded by the speaker, but is inferring the communicative intention of the speaker, in the context of the conversation. Grice's *cooperative principle* states that people must behave in a consistent and predictable manner to effectively communicate. Consistency is defined as adhering to maxims of quantity, quality, relation and manner. The terms of this contract create a rich pragmatic context for the task of understanding, which constrains how semantic, syntactic and lexical concerns are processed. Although Grice's work is a foundation of pragmatics, the maxims have proven quite difficult to formalize and apply in computational models (a fact which underscores the importance of the precision and thoroughness of computational modeling).

Clark and Haviland (H. H. Clark & Haviland, 1977) provided a notable refinement of Grice's co-operative contract with the proposal of a *given-new contract* between the speaker and the listener. They suggest that a key element of understanding is recognizing the given information separate from the new. The speaker is required to present information such that the hearer can always identify the proper antecedent for any given information, or be confident of the lack thereof. In many cases, such identification is indirect and requires *bridging inferences* from what is known or explicit to what is new. By casting a critical part of understanding in terms of a specific inferential task, Clark and Haviland began to move the intuitions of Grice's maxims towards a more formal, computational model. The given-new contract accounts for how the context of the discourse inherits from the context of the hearer, and how sentences are interpreted within that context.

5.1.2 Narrative as communication

These theories of communication are studied primarily in dialogue settings, where multiple agents have clear, often task-oriented, communication goals. However, the inferential model that they propose also has implications for narrative understanding. In classical speech act theory, narration is not considered an independent speech act, but a collection of assertion acts (Searle, 1975). It has been argued that this is not sufficient to account for either the obligations that surround a narration or the illocutionary force that it can exert (Nair, 2003). Being told a narrative is not merely a stand-in for directly observing the same set of events in the real world. Even a plain narration that lacks commentary or privileged insight must choose and enforce a point of view, selective attention, segmentation and ordering.

The distinction in narrative between what is being described and how it is presented is one of the fundamental elements of the literary field of narratology. Early Russian formalists took a more artifact-centric view of this distinction, treating the narrative as the result of the narration. Shklovsky and Tomashevski identified the *fabula* as a story as it would occur as a literal sequence of events, contrasted with the *sjuzhet*, a narrative that communicates that story (Bertens, 2008). Propp studied a hundred Russian folktales and concluded that they were all variations of the same story, different sjuzhets for the same fabula. In Propp's terminology, the fabula is a pattern of abstract character roles and functions (actions) to be filled and performed by characters, and each sjuzhet is a different way of instantiating them (Propp, 1971). Barthes, a French structuralist, proposed *functional units* as the smallest unit of narrative. In this view a narrative is not a static artifact but rather a sequence of dynamic acts. Some are *cardinal*, hinge-points of the plot, while others are *catalysers*, filling the story out with activity and diffuse concepts such as character and atmosphere (Barthes, 1977). Chatman later used the terms *story* and *discourse*, putting more emphasis on the narration process as a communicative act, in line with the pragmatics theories discussed above (Chatman, 1978).

The central assumption of structuralist narratology is that there are identifiable structures in narrative, and that those structures are what allow the narrative to be understood. If this is the case, then adhering to those structures follows Grice's cooperative principle. The reader, by identifying those expected structures, is led towards an understanding that meets the narrator's communication goals. While structuralism has been heavily criticized in literary studies as overly rigid and tied to the unfashionable idea that a narrative has a meaning of its own (apart from the reader), it is closer to a formal model of narrative understanding than any other school of criticism. Further, its weaknesses are far more significant when considering great works of literature and far less so when considering the simple stories of everyday communication.

5.1.3 Pragmatics and the meaning of a narrative

The American narratologist Gerald Prince directly addresses the question of how narratology accounts for the pragmatics of narrative (Prince, 1983). He argues that narrative understanding goes beyond grasping the content of the story and all its entailments within the story world. The act of telling the story, or writing it down, is an intentional communication that conveys some meaning. It could be a generally applicable lesson or a pointed statement that is highly specific to the context of the reader. Prince supports the view that relevance, in the Gricean sense, determines meaning for the reader. Relevance selects from among numerous coherent interpretations. To paraphrase Prince's examples, a story read by a Marxist critic could be interpreted as relating to the propositions of Marx, while for a student of Freud, Freud becomes a part of the story. A story could also be read in terms of enduring cultural values, current popular ideas or momentary preoccupations. Without pragmatic context, there is no way to account for the wide variety of possible meanings of a narrative.

This view accounts for the variety of possible meanings, and given the general reliability of conversational narrative communication, it predicts that shared context is quite accessible. However, relying on the definitions of relevance from dialogue research runs into a significant problem. Unlike in a dialogue situation, it is unlikely that the majority of the utterances in a narrative are personally relevant to the reader. This is especially true when considering a narrative about imaginary actors engaging in situations that never will or could occur. The events of the narrative are relevant to those actors, but not directly to the reader. It could be imagined that the reader is willing to wait until the end of the narrative for the relevant payoff, but that robs the theory of any usefulness in interpreting the narrative as it is taking place.

I argue that this problem provides evidence for the narratologists' claim that identifiable structures in the narrative guide understanding. These structures represent a shared context, which can be cooperatively used by the narrator and the reader. If they can act as a heuristic for relevance, they can lead the reader to infer an interpretation that meets the narrator's communication goals. For example, it may be that a bad thing happening to an imaginary character is assumed to be relevant to the reader, under the expectation that it will lead to events that really are relevant to the reader. The reader is following along with the conventions of narrative communication, which include the obligation of the narrator to provide a sufficiently relevant payoff.

Work by Labov followed a similar line of argument, proposing a narrative function of *evaluation* as a heuristic guide to meaning (Labov & Waletzky, 1966). Labov studied a series of personal, oral narratives where subjects talked about one of their experiences. He proposed a theory of three functional elements in the narratives, but found that even among those narratives that were well-formed, according to his theory, some were of much higher quality than others. This eventually led to the identification of a fourth function, evaluation, where the narrator would indicate that a certain clause or element of the story was more important than the rest. Labov showed that where such evaluations were used, they guided the hearer to a strong point of the story, while those narratives that lacked evaluation felt aimless and unsatisfactory. Much like Grice's maxims, Labov's evaluations have proven useful in analysis but difficult to formalize and apply.

5.1.4 Narrative structures

Propp's analysis of the structure of Russian folktales resulted in a set of 31 ordered functions, such as *Initial Situation*, *Villany* and *Victory*, parameterized by what character performs what role. Each of the folktales he studied is an instance of a subset of these functions, with the character roles assigned and the ordering maintained (Propp, 1971). Propp's work was very influential and gave strong evidence for common structure in narratives. However, the particular set of functions is very specific to Russian folktales. The complete, non-recursive ordering he specified strongly limits the range of narratives that can be expressed.

Work on memory and recall of narratives analyzed short stories to create hierarchical *story grammars*, such as that of Mandler and Johnson (Mandler & Johnson, 1977). The grammar was used to manually create story schemata, hypothesized as representative of human memory organization, to make predictions about what parts of stories would be remembered and recalled. Later work by Trabasso (Trabasso et al., 1984) created a simpler recursive transition network with six elements that could generate the structures of the more complex grammars in previous work. Trabasso's network reflects the key role of causality, demonstrated in his recall experiments. This network starts with a *setting (S)* that positions the protagonist in time and space, followed by one or more *initiating events (E)*. These events result in internal *reactions (R)* for the protagonist, which leads to at least one *goal state (G)*. The goal states provide motivation for *actions (A)* on the part of the protagonist which lead to *outcomes (O)* that do or do not satisfy the goals. The G-A-O pattern in the network can repeat, creating subsequent episodes, or recursively embed creating dependent sub-goals.

As tools for manual analysis, these theories of narrative structure provide insights into how typical narratives are constructed and the elements of narrative that people find most important and memorable. However, they have not been worked out to the level of precision necessary for a computational model of narrative understanding.

5.1.5 Inferring coherence and relevance

Work in computational models of narrative understanding has focused on inferring coherent interpretations. The early work by Charniak (Charniak, 1977) and Schank's group at Yale (Cullingford, 1978; Schank & Ableson, 1977; Wilensky, 1978) used frames, scripts, plans and other knowledge structures to identify events in a narrative as parts of larger events and situations. Using unfilled roles in these knowledge structures, their systems were able to infer implicit facts and set up expectations to aid disambiguation. They showed that narrative heavily uses assumed world knowledge, and that without it, deep understanding is not possible.

Dialogue-oriented work following from Grice's pragmatics has been concerned with the coherence of discourse structure in addition to coherence in the world being described. This has been modeled most effectively as *coherence relations* between pairs of utterances. Asher and Lascarides (Asher & Lascarides, 2003) and Hobbs (Hobbs, 1985) give formalized accounts of local coherence relations such as *explanation*, *elaboration* and *background* that connect sequential (nested) clauses. The constraint that each clause must have coherence relations with prior clauses guides semantic interpretation and can aid in disambiguation. Inferring these relations appeals to world knowledge, and both theories describe a logical interface to such knowledge. However, neither line of research has pursued implementation with a large-scale knowledge base.

Relevance Theory (RT) argues that relevance is a more effective alternative to coherence in driving the understanding process (Wilson & Sperber, 2004). Under this theory, understanding is constrained and directed by a search for relevance, defined in terms of *positive cognitive effects* – worthwhile updates to the hearer's world model. They demonstrate that relevance has a broader appeal to pragmatic and contextual factors than coherence (Wilson & Matsui, 2000) and argue that identifying specific coherence relations places an unnecessary burden on the process. Both Asher and Lascarides and Hobbs argue that their coherence relations are as effective as relevance as a guiding measure, but it may be more accurate to say that coherence relations implement a subset of what falls under relevance theory. However, those coherence-based formalisms have been worked out in much greater detail than relevance theory.

Judgments of relevance in RT are typically discussed in situated dialogue where the pragmatic concerns of the speaker and hearer may be invoked. If one is waiting for a train, then statements about the arrival time of that particular train are notably relevant and interpretation of ambiguous elements can be guided in that direction. However, within a narrative the notion of positive cognitive effect is insufficient to gauge the relevance of an utterance. The opening sentence in a story, for example "An ant went to a river to drink.", does not present any true conclusions nor does it reference any known entities or update any existing model. Instead, it establishes expectation. The imagined fact of an ant, being situated by a river and desiring a drink, will lead to further developments which the hearer can reasonably expect to be both coherent and relevant.

5.2 Narrative functions as a heuristic for relevance

I have argued that narrative structures can guide understanding as a heuristic for relevance. I hypothesize that certain classes of Barthes' cardinal functional units in narrative can serve as those structures. By attempting to interpret each utterance in a narrative as performing one or more of these functions, the interpretation process biases its understanding towards a coherent and relevant interpretation that meets the communicative goals of the narrator.

This view defines the interpretation process as the inferential task of fulfilling the expectations that certain narrative functions will be performed. This task is well suited for the practical EA NLU approach to language understanding. The expectations can be defined as a set of queries to be abductively proven by the discourse-level interpretation process. These proofs serve to disambiguate choice sets and generate an interpretation as a discourse-level DRS. Because this task is run on a sentence-by-sentence basis, disambiguation and discourse update happen incrementally. This allows the system to scale better with more complex reasoning tasks that would otherwise have to disambiguate large numbers of sentences in a single proof.

5.2.1 Expectations of narrative functions

Each narrative function defined here is represented by a predicate that relates a unit of the narrative to elements in the story. For example, if the first sentence in a story, *Sentence-3454063250-17665*, performs the function of introducing a new character, *ant17670*, then the following fact can be inferred:

(152) (introducesActor
(PresentationEventFn Sentence-3454063250-17665 IBTGeneration19746)
ant17670)

The concept of a unit of narrative is abstracted by the function *PresentationEventFn*, which accepts a sentence identifier and a uniquely generated event identifier as its arguments and denotes a presentation event. This event represents some presentation of elements in the narrative, within that sentence, that may be asserted to perform a particular narrative function. The sentence identifier alone is not sufficient to uniquely identify a presentation event, as there can be many events in any sentence. Likewise, a set of terms within the sentence is insufficient because the same words can fulfill more than one narrative function. Instead, the event identifier, *IBTGeneration19746* in this example, is used to uniquely identify a presentation event within a sentence. The expectation that this narrative function will appear is defined by the query form:

(153) (introducesActor (PresentationEventFn ?sentence-id ?event-id) ?actor)

The task of fulfilling this expectation is realized by attempting to abductively prove that there is some reasonable interpretation of the sentence (that is to say, consistent with the available choice sets and axioms) under which the query form is satisfied. This is accomplished using the EA NLU discourse interpretation facility described in section 2.4 and the assertion:

(154) (queryForInterpretation
0 (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))

together with the rule:

(155) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
(introducesActor (PresentationEventFn ?sentence-id ?event-id) ?actor))

The expectation that an actor might be introduced is always held and queried for in every sentence. In addition to those expectations that are always held, there are conditional expectations that bind to specific situations and entities in the ongoing narrative. These expectations are *opened* by another expectation being fulfilled and possibly *closed* when they are themselves fulfilled. They are queried for over the span of sentences when they are open. Expectations are opened by rules indicating that the fulfillment of one expectation results in the opening of another, with common bindings. For example, when a function presenting that an actor is aware of a situation is identified, it raises the expectation that a response by that actor to that situation is forthcoming. This is captured by the rule:

```
(156) (<== (openExpectation (PresentationEventFn ?sentence-id ?event-id)
                            (presentsResponse
                                (PresentationEventFn ?later-sentence-id ?other-event-id)
                                ?situation ?actor ?response)
           (presentsAwareness (PresentationEventFn ?sentence-id ?event-id)
                              ?actor ?situation))
```

According to this rule, when a presentation event is interpreted as performing a *presentsAwareness* function in the narrative, that presentation event is also asserted to open an expectation of a *presentsResponse* function. The bindings for the variables *?actor* and *?situation* in the *presentsResponse* query form are bound to their values in the *presentsAwareness* expression. All open expectations are queried for by the rule:

```
(157) (<== (meetsExpectation (PresentationEventFn ?sentence-id ?event-id)
                             (PresentationEventFn ?other-sentence-id ?other-event-id))
           (openExpectation (PresentationEventFn ?other-sentence-id ?other-event-id)
                            ?open-expectation)
          (uninferredSentence
            (closedExpectation (PresentationEventFn ?other-sentence2-id ?other-event2-id)
                               ?open-expectation))
           ?open-expectation)
```

where every open expectation query *?open-expectation* that is not closed will itself be queried. If that query succeeds, a *meetsExpectation* relation is inferred between the new, expected presentation event and the presentation event that opened the expectation.

Opening an expectation does not mean that it will be fulfilled. It means that the interpretation process will continue attempting to fulfill it. In this way, the set of open expectations guides the interpretation process towards certain readings of the narrative, which must be entailed by some possible disambiguation.

5.2.2 A theory of narrative functions

This theory is based on Barthes' notion of cardinal functional units, which drive the plot of a narrative, Trabasso's identification of narrative elements of causal importance and Labov's function of evaluation. The two patterns that make up Trabasso's networks, E-R-G and G-A-O, represent two fundamental structures in narrative: how actors respond to situations and how actors pursue goals. The results showing that these elements of causal networks are important in memory and recall (Trabasso et al., 1984) suggest that they perform a cardinal function.

One of the major limitations of Trabasso's networks, and story grammars in general, is that they assume a single point of view and do not account for interleaving patterns. A goal certainly may be expected to lead to actions undertaken to satisfy that goal, but numerous other goal introductions, goal-directed actions, and outcomes may occur in the meantime. The simpler part of this problem is the over-rigidity of the grammar. The more subtle part is that actions and events directed towards one goal may have an impact on another. In fact, that sort of interaction between goals is often an important part of what makes a story interesting. This argues for a larger set of relations than Trabasso presents, but it also suggests that a grammar approach, where constituents are uniquely categorized, is too limited. By contrast, the expectation fulfillment approach used here enforces only partial ordering, and allows a story element to participate in any number of functions.

The narrative functions defined here are named with either the prefix "presents" or the prefix "introduces" depending on whether the story element addressed by the function persists beyond its presentation.

5.2.2.1 Goals and goal-relevant situations

This theory defines goals in terms of a partial state of the world, expressed as a DRS. These logical models are valid or not at a certain time in the world described by the story. For example, the DRS for the character *ant17670* drowning is:

```
Universe: drown19751 ant17670

(isa drown19751 Drowning)
(bodilyDoer drown19751 ant17670)
```

Figure 24: DRS for ant17670 drowning

A goal is expressed as one of four functions of a DRS as follows:

(158) (Goal-AchieveFn (DrsCaseFn <drs-id>))

(159) (Goal-AlterFn (DrsCaseFn <drs-id>))

(160) (Goal-MaintainFn (DrsCaseFn <drs-id>))

(161) (Goal-AvoidFn (DrsCaseFn <drs-id>))

An achievement goal, shown in expression (158), is reached when the specified DRS becomes valid in the world of the story. An alteration goal, shown in expression (159), is reached when the specified DRS becomes invalid. A maintenance goal, shown in expression (160), is maintained as long as the specified DRS remains valid. An avoidance goal, shown in expression (161), is maintained as long as the specified DRS remains invalid. Goals are held by actors, but a particular goal is not reified specific to a particular actor. This is important because a goal can be held by several actors, by different actors at different times or by no actors and spoken of only hypothetically.

The narrative function *introducesGoal* communicates that an actor holds a goal. The expectation that goals will be introduced is always active, as implemented by the rule:

(162) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
(introducesGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?goal)

In general, goals are introduced through explicit statements of desire or intention, or through commonsense reasoning such as, for example, an explicit statement of hunger implies a goal of eating. Universal goals such as surviving or avoiding pain are also introduced into the narrative by their violation.

The introduction of a goal in the narrative opens several expectations. First, there is the expectation that subsequent situations in the narrative should be evaluated with respect to this goal. This is captured by the rule:

(163) (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
(evaluatesOutcome
(PresentationEventFn ?later-sentence-id ?other-event-id)
?situation ?goal ?outcome))
(introducesGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?goal))

The idea of evaluating a situation is importantly distinct from any notion of how that situation came about. It could be that an actor holding the goal acted intentionally to bring about a state which satisfied the goal. Or it could be that an action by another actor, directed towards the goal or not, or even a non-agentive event resulted in the goal succeeding or failing. The set of possible outcomes, which are valid to bind to the variable *?outcome*, are *Success*, *Failure*, *Success-Partial* and *Failure-Partial*. Success and failure are judged differently for the four different types of goals. An achievement or alteration goal succeeds when the world is changed such that the DRS of the goal becomes valid or invalid, respectively. Such a goal fails only when it is no longer possible for it to succeed. A maintenance or avoidance goal, on the other hand, fails when the world is changed such that the DRS of the goal becomes invalid or valid, respectively. These goals succeed only when it is no longer possible for them to fail. The -*Partial* categories account for cases where a goal may be incrementally satisfied or violated, such as when multiple entities are involved in the goal state. An evaluation of *Success* or *Failure* also closes all expectations related to that goal.

Often the cause of a goal-relevant outcome is also relevant. This is captured by the narrative function *presentsResult*. While it is possible that results are of general interest, they are at this point limited to results that are also goal-relevant outcomes. Thus they are not expected, but rather queried for fulfilled expectation outcomes. This is captured by the rule:

```
(164) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
           (evaluatesOutcome (PresentationEventFn ?sentence-id ?other-event-id)
                             ?result ?goal ?outcome))
           (presentsResult (PresentationEventFn ?sentence-id ?event-id) ?cause ?result))
```

Goals are tightly linked to plans intended to realize them. Plans connect actions to the goals they are intended to satisfy, express choices between actions that might be undertaken, and create hierarchical goal dependencies. Plan reasoning is a key part of explaining character actions, but this theory of narrative functions does not reify plans as an element of interest like goals. Rather, it expresses the relations between goals, actions and sub-goals as direct expectations.

A sub-goal is asserted where an actor believes that one satisfying goal is a possible step in satisfying another. A sub-goal may stem from a complete, pre-conceived plan or from a complex interaction of plans and changing circumstances. When a goal is introduced, possible sub-goals functions are expected as captured by the rule:

```
(165) (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
                    (subGoal
                        (PresentationEventFn ?later-sentence-id ?other-event-id)
                        ?actor ?sub-goal ?goal))
          (introducesGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?goal))
```

This is a recursive relation, such that when a sub-goal is identified it also introduces itself as a goal:

```
(166) (<== (introducesGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?sub-goal)
           (subGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?sub-goal ?goal))
```

A goal-directed action is an action undertaken by an actor who holds a goal and believes the action is a possible step in satisfying it. The expectation that such actions will follow the introduction of a goal is captured by the rule:

```
(167) (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
                             (presentsAction
                                (PresentationEventFn ?later-sentence-id ?other-event-id)
                                 ?actor ?action ?goal)
           (introducesGoal (PresentationEventFn ?sentence-id ?event-id) ?actor ?goal))
```

Unlike earlier theories, this theory goes beyond goal-directed plan actions to account for three classes of interactions with other agents' actions and circumstances. The first two classes are threats and obstacles. A threat occurs when an action or event raises the possibility of causing a goal to fail. In the case of a threatening action, it is not necessary that the agent responsible for the action intend to threaten the goal or even be aware of it. An obstacle, by contrast, is a situation that must be resolved before a goal can succeed. The third class of interaction involves an opportunity: a situation that enables some plan of action to see the goal succeed. These expectations are raised when a goal is introduced and captured by the rules:

```
(168)  (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
                              (introducesThreat
                                 (PresentationEventFn ?later-sentence-id ?other-event-id)
                                 ?event ?goal)
            (introducesGoal (PresentationEventFn ?sentence-id ?event-id)?actor ?goal))
```

(169) (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
(introducesObstacle
(PresentationEventFn ?later-sentence-id ?other-event-id)
?situation ?goal)
(introducesGoal (PresentationEventFn ?sentence-id ?event-id)?actor ?goal))

(170) (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
(introducesOpportunity
(PresentationEventFn ?later-sentence-id ?other-event-id)
?situation ?goal)
(introducesGoal (PresentationEventFn ?sentence-id ?event-id)?actor ?goal))

5.2.2.2 Non-goal-directed responses

Not all situations and events in a narrative are goal-directed. Some might impact any number of goals were the story world to be realized in complete detail, but those goals are not relevant to the narrative and do not need to be identified. Trabasso's causal networks account for non-goal-directed events of interest through the E-R-G pattern, recognizing that narratives may start with an event whose causes do not matter, but which leads to a relevant goal. More generally, a narrative may involve actions whose specific motivations are unimportant, but nevertheless cause a chain of behaviors and outcomes.

Some narratives begin with a background activity that contributes to setting and atmosphere rather than plot. The expectation that such activity will be presented is always active, and implemented by the rule:

(171) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
(presentsActivity (PresentationEventFn ?sentence-id ?event-id) ?actor ?event))

In the absence of goals as a focusing mechanism, I suggest that awareness can serve a similar purpose. The narrative function of presenting awareness either shows that an actor has become aware of a situation or explicitly points out that an actor is aware of a situation. Awareness here is not knowledge, it could be something that they already knew and are merely said to be focusing on. The expectation that awareness will be presented is always active, and implemented by the rule:

```
(172)  (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
            (presentsAwareness (PresentationEventFn ?sentence-id ?event-id)
                               ?actor ?situation))
```

Presenting awareness of a situation, which might otherwise have no significance, begs the question of why it matters and will the aware actor deliberately respond in some way? When awareness is presented, it raises the expectation of a response. Unlike goal-directed action, this response is not tied to any goal that has been presented as significant in the narrative. The motivation of the actor may become clear later, or may not be important, and the burden is on the narrator to make those distinctions clear. This is captured by the rule:

```
(173)  (<== (opensExpectation (PresentationEventFn ?sentence-id ?event-id)
                              (presentsResponse
                                (PresentationEventFn ?later-sentence-id ?other-event-id)
                                 ?situation ?actor ?response))
            (presentsAwareness (PresentationEventFn ?sentence-id ?event-id)
                               ?actor ?situation))
```

Two notable sub-cases of awareness are being addressed in a dialogue and being aware of a goal succeeding or failing. In both cases a response is expected.

5.2.2.3 Labovian evaluation

Two evaluation functions are included here: *presentsSymmetry* and *presentsContrast*. Unlike the rest of the functions, these are not cardinal functional units that advance the story. Rather, they are functions of the structure of the presentation itself. Symmetry is identified when multiple sequences of actions or events repeat or parallel each other. Symmetry serves to emphasize a repeated action or situation or to highlight contrast between them. Contrast is identified when two story elements are set up against each other. This may be accomplished through symmetry, through an explicitly stated relationship such as comparison, or through the use of an explicit utterance describing alternatives to the events in the world (or in another utterance). The expectation of these evaluations is always active, implemented by the rules:

(174) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
(presentsSymmetry (PresentationEventFn ?sentence-id ?event-id)
?sit-or-prop1 ?sit-or-prop2))

(175) (<== (narrativeFunction (PresentationEventFn ?sentence-id ?event-id))
(presentsContrast (PresentationEventFn ?sentence-id ?event-id)
?sit-or-prop1 ?sit-or-prop2))

Each function presents two story elements, either situations or propositions. The propositions are facts that are true in the world of the story.

5.2.3 Abducing narrative functions

Fulfilling these expectations is carried out by a set of queries in *EANarrativeQueriesMt*. An abductive query using this microtheory attempts to prove the presence of all the expected functions in the same proof context. It is possible that this will result in multiple interpretations, with conflicting assumption sets. In this case, the interpretation that maximizes the *relevance* *score* is chosen. The relevance score for an interpretation is the sum of the scores for each narrative function entailed by that interpretation. The relevance score for a narrative function is calculated using the recursive algorithm given in Figure 25. Given a narrative function *nfn*, the algorithm collects all prior functions *nfn'* for which it is true that *(meetsExpectation nfn nfn')*. The score for *nfn* is a constant *C*, multiplied by the scores of all the collected functions *nfn'* (unless there are none). For example, there is always an active expectation for the function *introducesGoal*, so it never *meetsExpectation* for a prior function. When a goal is introduced, it always has a score of *C*. Introducing a goal also opens an expectation for *evaluatesOutcome*, among others. If a later sentence presents a function that meets that expectation, that function will have a score of *C* multiplied by the score of the *introducesGoal* function. The score for the *evaluatesOutcome* will therefore by C*C. Higher scores are possible when a function meets the expectations of multiple prior functions. This scoring algorithm biases the system towards chains of specifically expected functions.

```
relevance( nfn ) =
gather all nfn' such that:
  (meetsExpectation nfn nfn')
if nfn' = {} then
  C
else
  C * ∏ relevance(nfn')
```

Figure 25: Recursive algorithm for calculating the relevance of a narrative function *nfn*

Because the *meetsExpectation* relation is unidirectional from earlier to later arguments, there are no cycles to be concerned with. A procedure that computes the relevance score is passed in to the abductive query as the measure of the quality of the explanation as described in section 4.3.1.

5.3 Understanding fables

Fables are short stories characterized by anthropomorphized animals, objects and forces that illustrate a moral lesson. These stories exist in literature worldwide, often originating in ancient times as part of an oral storytelling tradition. For example, Aesop's fables were published in England in the 17th century and enjoyed immense popularity, solidifying the fable as a distinct genre in English literature. These fables are attributed to the Greek slave Aesop, believed to have lived around 550BC (Lewis, 1996).

Aesop's fables exemplify the use of narrative to convey a specific point. The inclusion of a *moral*, a one-sentence maxim separate from the story proper, strongly indicates authorial intent. The reading of a fable is still subjective, and a given fable may support several different morals, or even poorly support its own moral, but the popularity and endurance of many of Aesop's fables suggests that they do effectively communicate their point. These fables provide a more specific instance of the general problem of identifying possible meanings of a narrative. Reading the story part of the fable should create a discourse context in which the moral can be understood as both a restatement and an expansion of the point being illustrated.

This evaluation tests the ability of EA NLU to identify which of a set of morals best fits with a fable story. There are two goals in this evaluation. The first is to test the sufficiency of the EA NLU semantic interpretation process and the representations it generates for this reasoning task. The second is to test the sufficiency of the theory of narrative expectations presented in this chapter to guide that interpretation process.

5.3.1 Selecting fables

The web site aesopfables.com, available since 1996, hosts a publically available collection of 638 fables (and growing). The site contains a set of 86 “Selected Fables”, chosen by the maintainer for “ease of reading and concise moral understanding”. The majority of the fables, including all the selected fables, were translated into English in the eighteenth century by Rev. George Fyler Townsend.

In order to select a diverse cross-section of fables for this task, I categorized the selected fables according to both the surface form of the story proper and the surface form of the moral. The stories broke down into two major categories, while the morals presented some clear categories and others less so.

The first story form presents an outcome. The story concludes with an event that is the result of what went before. This event must carry a positive or negative valence for the characters. This valence can be utilitarian or moral and can be based on goals and values identified in the story context or generally accepted in the context of the storytelling. This type of story bases its argument on a depiction of causality, implying that there is some level of generality to the good or bad thing that happened as a result of character, action or setting. 29 of the selected fables are of this type.

The second story form observed in the selected fables is an explicit utterance that states a particular insight regarding the story situation. The utterance may take the form of a rhetorical question, observation or answer, but the content is always a reflective interpretation of the situation rather than a character-specific response. As a simple rule of thumb, these insights could be spoken by any character, or the narrator, with essentially the same effect. In many cases they express an explanation for what, how or why something did, will or should have happened. In others they assert a matter-of-fact truth about the situation. 37 of the selected fables are of this type.

14 of the selected fables present the moral as the last line of dialogue in the story. Consequently, they are not suitable for a task requiring the system to identify a detached moral. The remaining 6 fables are similar to the outcome form, but conclude with an action that is not positive or negative. Instead, they conclude with an action on the part of the protagonist in response to some state or event. The response places a particular emphasis on the causal relation leading to action and implies some level of generality about how people respond to situations. While the response may be verbal, it is distinct from the reflective commentaries that mark the insight form. This type of utterance is specific to the character that performs it and continues the causal chain of the story.

There are three moral forms that can be clearly categorized in the selected fables. 11 of the morals are imperative statements giving direct advice to the hearer. They either exhort or discourage a certain action. 15 of the morals make a judgment by assigning a value to a situation or action based on some, often underspecified, scale. This can be accomplished by comparing the relative value of two elements or by relating a single element to a generic quantity. 9 of the morals present a revelation regarding two contrasting situations. They suggest the possibility that the one situation leads to or can be viewed as the other.

This evaluation uses these categorizations to select a set of fables that are reasonably diverse. For each of the two major categories of story forms, one fable with each of the recognized moral forms is selected. Table 11 lists the six fables used.

Fable	**Story Form**	**Moral Form**
The Ass, the Fox and the Lion	Outcome	Advice
The Cat and Venus	Outcome	Judgment
The Dove and the Ant	Outcome	Revelation
The Boy and the Nettles	Insight	Advice
The Dogs and the Fox	Insight	Judgment
The Boys and the Frogs	Insight	Revelation

Table 11: Selected fables

5.3.2 Extending EA NLU grammar and semantic frames

QRG-CE has been extended for each set of narratives described in this dissertation, as well as for non-narrative text used in several other projects. One of the key questions for the practical EA NLU approach is how many extensions must be made for each new narrative, and whether that effort is constant or decreasing as more narratives are processed. Table 12 summarizes the grammar extensions made to process each of the fables. For example, in *The Ass, the Fox and the Lion*, 38 of the rules in QRG-CE are used at least once in at least one of the selected parse trees. The number of those grammar rules which were added to QRG-CE in order to parse each fable is also given. These numbers are obviously dependent on the variable length of the stories, so the rule extension is given as a percentage as well. Figure 26 shows those percentages as a bar chart.

Fable	Grammar Rules Used	Grammar Rules Added	Percent New Rules
The Dove and the Ant	25	5	20%
The Ass, the Fox and the Lion	38	6	16%
The Dogs and the Fox	23	2	9%
The Boy and the Nettles	19	1	5%
The Boys and the Frogs	33	1	3%
The Cat and Venus	17	1	6%

Table 12: Grammar use and extension for the fables

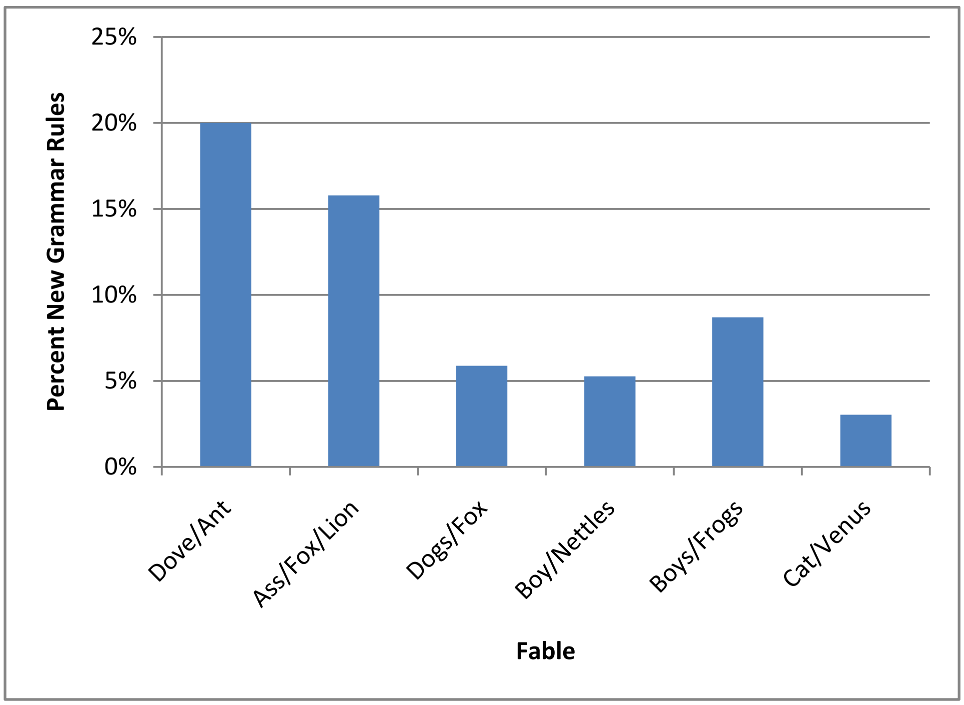

Figure 26: Percent of new grammar rules used to parse the fables

What is most notable about this result is that the first two fables, *The Dove and the Ant* and *The Ass, the Fox and the Lion* were parsed several months before the other four fables. In the intervening time, numerous updates were made in support of other projects and experiments, including the Iranian folktales discussed in Chapter 2. The current version of the grammar, used to parse the last four fables, required very little extension to handle them. This is an encouraging sign that the grammar coverage is being improved in general by project-these specific extensions.

Table 13 and Figure 27 show the same data for the semantic frames used and added to the knowledge base in the processing of each fable. Again, these are the frames that are selected by the semantic interpretation process in the moral matching task. Given the diversity of domains represented in the fables, it is not surprising that the number of semantic frames required is fairly random. As discussed in chapter 2 and chapter 4, adding semantic frames is a low effort extension, and the semantic interpretation process scales well with additional frames.

Fable	**Semantic Frames Used**	**Semantic Frames Added**	**Percent New Frames**
The Dove and the Ant	38	10	26%
The Ass, the Fox and the Lion	45	12	27%
The Dogs and the Fox	20	3	15%
The Boy and the Nettles	15	8	53%
The Boys and the Frogs	15	4	27%
The Cat and Venus	35	16	46%

Table 13: Semantic frame use and extension for the fables

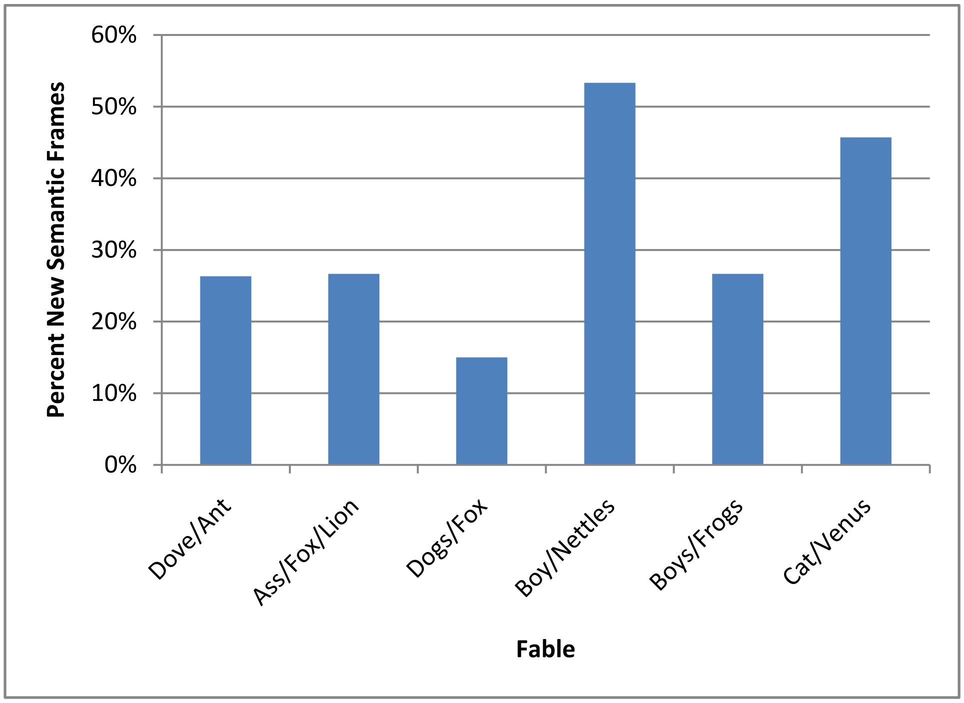

Figure 27: Percent of new grammar rules used to parse the fables

5.3.3 Narrative expectations and interpretation

The theory of narrative expectations implemented by *EANarrativeQueriesMt* is used as the pragmatic context for interpreting the fable stories. As an overview, Table 14 shows the narrative functions defined in the theory and how many times they are inferred in the fables. The total instances of each function and the number of fables they are found in (as a count and a percent) are given separately for the three outcome fables and the three insight fables. The outcome fables are *The Dove and the Ant, The Ass, the Fox and the Lion* and *The Cat and Venus*. The insight fables are *The Dogs and the Fox, The Boy and the Nettles* and *The Boys and the Frogs*.

Narrative functions inferred in outcome fables			
Narrative function	**Instances**	**Used in Fables**	**Used in Percent of Fables**
introducesActor	10	3	100%
introducesGoal	13	3	100%
presentsAction	9	3	100%
introducesThreat	3	2	67%
introducesOpportunity	2	2	67%
introducesObstacle	1	1	33%
presentsResult	2	1	33%
evaluatesOutcome	12	3	100%
presentsActivity	0	0	0%
presentsAwareness	6	3	100%
presentsResponse	8	3	100%
presentsContrast	2	2	67%
presentsSymmetry	1	1	33%
Narrative functions inferred in insight fables			
Narrative function	**Instances**	**Used in Fables**	**Used in Percent of Fables**
introducesActor	7	3	100%
introducesGoal	3	3	100%
presentsAction	1	1	33%
introducesThreat	0	0	0%
introducesOpportunity	0	0	0%
introducesObstacle	0	0	0%
presentsResult	1	1	33%
evaluatesOutcome	3	3	100%
presentsActivity	1	1	33%
presentsAwareness	5	3	100%
presentsResponse	7	3	100%
presentsContrast	3	3	100%
presentsSymmetry	0	0	0%

Table 14: Narrative functions in the outcome fables

Introducing actors is naturally a consistent function in all the fables. Likewise, evaluating outcomes is common in all the fables, with some explicit causal result links leading to those outcomes. Presentation of awareness and response is also widely used. The main difference lies in that the insight fables state only implicit goals of not being hurt or killed, and there is little goal-directed action to move the plot along. Those stories instead rely mostly on presenting awareness and non-goal-directed responses. Contrast is found in all but one of the fables which instead presents symmetry. The next sections contain detailed analysis of the interpretation of each fable.

5.3.3.1 The Ass, the Fox and the Lion

THE ASS and the Fox, having entered into partnership together for their mutual protection, went out into the forest to hunt. They had not proceeded far when they met a Lion. The Fox, seeing imminent danger, approached the Lion and promised to contrive for him the capture of the Ass if the Lion would pledge his word not to harm the Fox. Then, upon assuring the Ass that he would not be injured, the Fox led him to a deep pit and arranged that he should fall into it. The Lion, seeing that the Ass was secured, immediately clutched the Fox, and attacked the Ass at his leisure.

Never trust your enemy.

Figure 28: The Ass, the Fox and the Lion

An Ass and a Fox, having entered into a partnership for protection, went into the forest to hunt.

They had proceeded a short distance when they met a Lion.

The Fox, seeing danger, approached the Lion and promised to capture the Ass for him.

In return, the Lion would promise to not harm the Fox.

The Fox, having assured the Ass that he would not be injured, led him to a deep pit.

He then caused the Ass to fall into it.

The Lion, seeing that the Ass was trapped, immediately attacked the Fox.

He then attacked the Ass at his leisure.

Do not trust your enemy.

Figure 29: QRG-CE version of The Ass, the Fox and the Lion

Figure 28 contains the original text for the story *The Ass, the Fox and the Lion* and Figure 29 contains the version rendered in QRG-CE. This version divides up particularly complex clauses and uses a few simpler rewordings. A notable change is that the negative assertion in the second sentence is replaced with a positive form that is more easily represented, some more straightforward terms (e.g. “trapped”, “attacked”) are used and the universal quantifier “never” in the moral is replaced with a direct statement. Due to our focus on scenarios and other concrete narratives, EA NLU currently has limited support for generic statements.

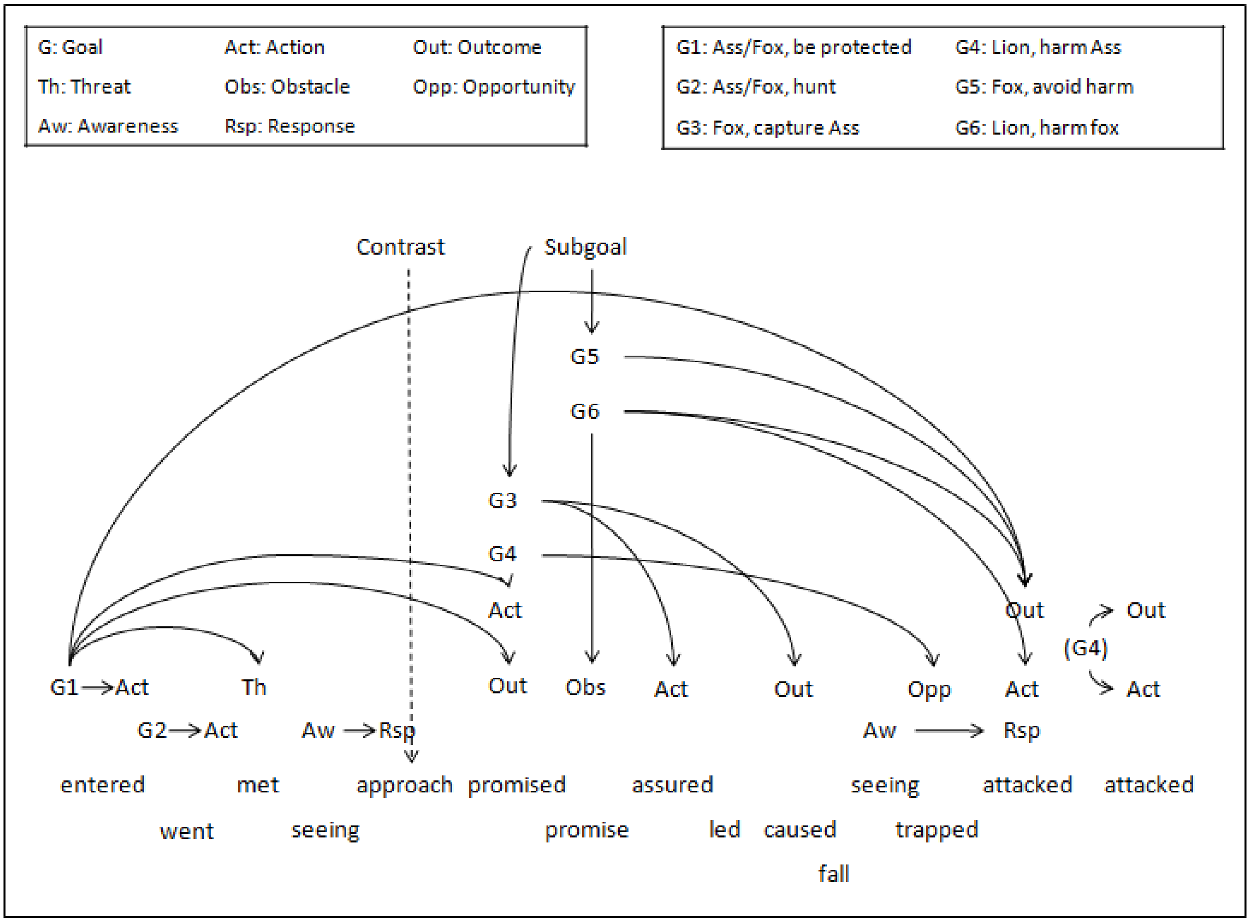

Figure 30: Narrative functions for The Ass, the Fox and the Lion

Figure 30 shows the narrative functions inferred in the interpretation of *The Ass, the Fox and the Lion* and the story events that they anchor to. Functions are placed over the event whose presentation they represent, with additional dashed arrows added for evaluative functions. Note that the fourth goal *G4* is repeated in parenthesis before the final event in the story to eliminate some clutter. The function *introducesActor*, which does not link to other functions, is omitted to keep the diagram more readable. This section contains a sentence-by-sentence analysis of this interpretation to ensure that the details are clear. The narrative functions inferred are given in their entirety. The meta-relations *opensExpectation*, *meetsExpectation* and *closedExpectation* are not included here for readability. The following sections for the other fables are abbreviated, with their complete details, including meta-relations, given in Appendix C.

An Ass and a Fox, having entered into a partnership for protection, went into the forest to hunt.

This narrative begins by introducing two actors, the Fox and the Ass, who each have the same goal of being protected. These goals have led to the goal-directed action of entering into a partnership together. Having done so, they are now pursuing a common goal of hunting, for which they have gone into the forest. The narrative functions inferred are:

(176) (introducesActor
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24712)
fox21054)

(177) (introducesActor
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24713)
ass21039)

(178) (introducesGoal
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24714)
fox21054 (Goal-MaintainFn protection21295))

(179) (introducesGoal
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24715)
ass21039 (Goal-MaintainFn protection21295))

(180) (introducesGoal
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24716)
fox21054 (Goal-AchieveFn (DrsCaseFn DRS-3454235456-24716)))

(181) (introducesGoal
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24717)
ass21039 (Goal-AchieveFn (DrsCaseFn DRS-3454235456-24716)))

(182) (presentsAction
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24710)
fox21054 go21456 (Goal-AchieveFn (DrsCaseFn DRS-3454235456-24716)))

(183) (presentsAction
(PresentationEventFn Sentence-3454235422-21034 IBTGeneration24711)
ass21039 go21456 (Goal-AchieveFn (DrsCaseFn DRS-3454235456-24716)))

where the DRS in the goal statements, *DRS-3454235456-24716*, describes the partners hunting. Inferring these narrative functions abductively disambiguates choice sets in the first sentence that entail this interpretation. Not all choice sets are disambiguated. The fragment of the discourse-level DRS for this sentence, omitting the universe of the top-level DRS, is shown in Figure 31.

```
Universe: ass21039 protection21295 fox21054 forest21551 go21456 partnership21209 enter21130

(isa ass21039 Donkey)
(isa fox21054 Fox)

(isa enter21130 BecomingAParticipantInSomething)
(performedBy enter21130 fox21054)
(performedBy enter21130 ass21039)
(outputsCreated enter21130 partnership21209)

(purposeInEvent fox21054 enter21130 protection21295)
(purposeInEvent ass21039 enter21130 protection21295)

(isa partnership21209 Partnership)
(socialParticipants partnership21209 fox21054)
(socialParticipants partnership21209 ass21039)

(isa protection21295 ProtectionSituation)
(thingProtected-Generic protection21295 ass21039)
(thingProtected-Generic protection21295 fox21054)

(after go21456 enter21130)

(isa go21456 Movement-TranslationEvent)
(primaryObjectMoving go21456 fox21054)
(primaryObjectMoving go21456 ass21039)

(purposeInEvent ass21039 go21456 (DrsCaseFn DRS-3454235456-24716))
(purposeInEvent fox21054 go21456 (DrsCaseFn DRS-3454235456-24716))

DRS-3454235456-24716:
    Universe: hunt21636

    (isa hunt21636 Hunting)
    (performedBy hunt21636 fox21054)
    (performedBy hunt21636 ass21039)
```

Figure 31: Partial DRS for the first sentence of The Ass, the Fox and the Lion

They had proceeded a short distance when they met a Lion.

The second sentence introduces the lion as an actor and as a threat to the protection goal held by the partnership. The narrative functions inferred are:

(184) (introducesActor
(PresentationEventFn Sentence-3454235425-21792 IBTGeneration24717)
lion22287)

(185) (introducesThreat
(PresentationEventFn Sentence-3454235425-21792 IBTGeneration24723)
Situation24718 (Goal-MaintainFn protection21295))

The relevant fragment of the discourse-level DRS is shown in Figure 32.

```
(isa proceed21854 GoingSomewhere)

(after meet22160 proceed21854)

(isa meet22160 EncounteringSomething)
(doneBy meet22160 lion22287)
(encounteredObject meet22160 partnership21209)

(isa lion22287 Lion)

(holdsIn Situation24718 (near partnership21209 lion22287))
```

Figure 32: Partial DRS for the second sentence of The Ass, the Fox and the Lion

The Fox, seeing danger, approached the Lion and promised to capture the Ass for him.

The third sentence presents the awareness of the fox of the dangerous situation at hand and his following actions are interpreted as a response to that awareness as well as a goal-directed action towards being protected. The promise made by the fox to the lion introduces a goal on the fox's part to capture the ass and a goal on the lion's part to harm the ass. It also partially fails the partnership's initial goal that they both be protected. The presentation of contrast is inferred, based on the explicit approach undertaken by the fox, between the awareness of the fox and the ass regarding the promise being made. The betrayal is not explicit at this level, being more specialized than the theory deems necessary. However, it does capture an interesting pattern where the fox deliberately acts to forward the goal of being protected (for himself), directly causing the partial failure of that goal (for the ass). The narrative functions inferred are:

(186) (presentsAwareness
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24738)
fox21054 Situation24718)

(187) (presentsResponse
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24739)
Situation24718 fox21054 PurposefulAction24737)

(188) (presentsAction
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24734)
fox21054 PurposefulAction24737 (Goal-MaintainFn protection21295))

(189) (introducesGoal
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24735)
fox21054 (Goal-AchieveFn (DrsCaseFn DRS24732)))

(190) (introducesGoal
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24736)
lion22287 (Goal-AchieveFn (DrsCaseFn DRS24733)))

(191) (evalutatesOutcome
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24732)
PurposefulAction24737 (Goal-MaintainFn protection21295) Failure-Partial)

(192) (presentsContrast

(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24740)
(aware fox21054 promise22543) (not (aware ass21039 promise22543)))

where *PurposefulAction24737* is an action made up of the approaching and promising actions, the fox's goal DRS, *DRS24732*, describes the fox capturing the ass and the implicit goal DRS for the lion, *DRS24733*, describes his harming the ass. The relevant fragment of the discourse-level DRS is shown in Figure 33. The implicit goal DRS for the lion is shown in Figure 34.

(isa see22415 VisualPerception)
(performedBy see22415 fox21054)
(perceivedThings see22415 danger22437)

(isa danger22437 DangerousSituation)

(isa approach22463 Approaching)
(primaryObjectMoving approach22463 fox21054)
(toLocation approach22463 lion22287)

(isa promise22543 MakingAPromise)
(senderOfInfo promise22543 fox21054)
(recipientOfInfo promise22543 lion22287)
(promiseStatement promise22543 (DrsCaseFn DRS24732))

DRS24732:

Universe: lion22287 ass21039 capture22608

(isa ass21039 Donkey)

(isa capture22608 CapturingSomething)
(performedBy capture22608 fox21054)
(objectActedOn capture22608 ass21039)

(subEvents PurposefulAction24737 approach22463)
(subEvents PurposefulAction24737 promise22543)

Figure 33: Partial DRS for the third sentence of The Ass, the Fox and the Lion

Universe: HarmingAnAgent24734

(isa HarmingAnAgent24734 HarmingAnAgent)
(doneBy HarmingAnAgent24734 lion22287)
(maleficiary HarmingAnAgent24734 ass21039)

Figure 34: DRS for the lion harming the ass

In return, the Lion would promise to not harm the Fox.

The fourth sentence presents another implicit goal, this time that the lion would also like to harm the fox (thus the need to prohibit it), and a goal for the fox of avoiding that harm. The promise being required of the lion is interpreted as presenting an obstacle to that implicit goal. Meanwhile, the fox's promise to capture the ass is asserted as a sub-goal of his goal of avoiding harm. The narrative functions inferred are:

(193) (introducesGoal
(PresentationEventFn Sentence-3454235432-22958 IBTGeneration24746)
lion22287 (Goal-AchieveFn (DrsCaseFn DRS24743)))

(194) (presentsObstacle
(PresentationEventFn Sentence-3454235432-22958 IBTGeneration24745)
promise23029 (Goal-AchieveFn (DrsCaseFn DRS24743)))

(195) (introducesGoal
(PresentationEventFn Sentence-3454235432-22958 IBTGeneration24748)
fox21054 (Goal-AvoidFn (DrsCaseFn DRS24743)))

(196) (presentsSubgoal
(PresentationEventFn Sentence-3454235432-22958 IBTGeneration24749)
fox21054
(Goal-AchieveFn (DrsCaseFn DRS24732)
(Goal-AvoidFn (DrsCaseFn DRS24743))

where the DRS for the new goals, *DRS24743*, describes the lion harming the fox and the sub-goal DRS, *DRS24732*, describes the fox capturing the ass (shown in Figure 33). The relevant fragment of the discourse-level DRS is shown in Figure 35.

```
(isa return22981 MakingAnAgreement)
(agreeingAgents return22981 lion22287)
(requestStatement return22981 (DrsCaseFn DRS-3454235641-24749))

DRS-3454235641-24749:
  Universe: lion22287

  (possible-Historical (DrsCaseFn DRS-3454235641-24750))

  DRS-3454235641-24750:
    Universe: promise23029

    (isa promise23029 MakingAPromise)
    (senderOfInfo promise23029 lion22287)
    (promiseStatement promise23029 (not (DrsCaseFn DRS24743)))

    (isa lion22287 Lion)

    DRS24743:
      Universe: fox21054 harm23081

      (isa harm23081 HarmingAnAgent)
      (doneBy harm23081 lion22287)
      (maleficiary harm23081 fox21054)

      (isa fox21054 Fox)
```

Figure 35: Partial DRS for the fourth sentence of The Ass, the Fox and the Lion

The Fox, having assured the Ass that he would not be injured, led him to a deep pit.

The fifth sentence presents actions undertaken by the fox which are interpreted as a goal-directed action aimed at capturing the ass. The narrative function inferred is:

(197) (presentsAction
(PresentationEventFn Sentence-3454235433-23229 IBTGeneration24758)
fox21054 PurposefulAction24757 (Goal-AchieveFn (DrsCaseFn DRS24732)))

where *PurposefulAction24757* is an action made up of the assuring and leading actions and the DRS for the goal, *DRS24743*, describes the fox capturing the ass (shown in Figure 33). The relevant fragment of the discourse-level DRS is shown in Figure 36.

(isa assure23296 Reassurance)
(performedBy assure23296 fox21054)
(agentInfluenced assure23296 ass21039)
(infoTransferred assure23296 (DrsCaseFn DRS-3454235667-24759))

DRS-3454235667-24759:

> Universe: ass21039
>
> (not (possible-Historical (DrsCaseFn DRS-3454235667-24760)))
>
> DRS-3454235667-24760:
>
> > Universe: be23408
> >
> > (hasPhysiologicalFeature ass21039 Injured)

(after lead23676 assure23296)

(isa lead23676 GuidingAMovingObject)
(primaryObjectMoving lead23676 ass21039)
(directingAgent lead23676 fox21054)

(isa pit23800 Pit-Topographical)

(subEvents PurposefulAction24757 assure23296)
(subEvents PurposefulAction24757 lead23676)

Figure 36: Partial DRS for the fifth sentence of The Ass, the Fox and the Lion

He then caused the Ass to fall into it.

The sixth sentence presents the successful outcome of the fox's goal of capturing the ass. The narrative function inferred is:

(198) (evaluatesOutcome
(PresentationEventFn Sentence-3454235436-23959 IBTGeneration24768)
PurposefulAction24767 (Goal-AchieveFn (DrsCaseFn DRS24732)) Success)

where *PurposefulAction24767* is the implicit action in the sentence that caused the ass to fall and the DRS for the goal, *DRS24743*, describes the fox capturing the ass (shown in Figure 33). The relevant fragment of the discourse-level DRS is shown in Figure 37.

```
(performedBy PurposefulAction24767 fox21054)
(causes-SitProp PurposefulAction24767 (DrsCaseFn DRS24766))

DRS24766:
  Universe: ass21039 pit23800 fall24068

  (isa fall24068 FallingEvent)
  (primaryObjectMoving fall24068 ass21039)
  (into-UnderspecifiedContainer fall24068 pit23800)

  (isa ass21039 Donkey)
```

Figure 37: Partial DRS for the sixth sentence of The Ass, the Fox and the Lion

The Lion, seeing that the Ass was trapped, immediately attacked the Fox.

The seventh sentence presents the awareness of the lion that the ass has been captured, which represents an opportunity for him to fulfill his assumed goal of harming the ass. He responds to that awareness with the goal-directed action of attacking the fox, thus fulfilling his other goal and simultaneously failing the fox's goal of not being harmed and the original goal of being protected. The narrative functions inferred are:

(199) (presentsAwareness
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247673)
lion22287 trap24296)

(200) (presentsResponse
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247674)
trap24296 lion22287 attack24383)

(201) (introducesOpportunity
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247675)
trap24296 (Goal-AchieveFn (DrsCaseFn DRS24733)))

(202) (presentsAction
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247677)
lion22287 attack24383 (Goal-AchieveFn (DrsCaseFn DRS24743)))

(203) (evaluatesOutcome
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247678)
attack24383 (Goal-AchieveFn (DrsCaseFn DRS24743)) Success)

(204) (evaluatesOutcome
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration247679)
attack24383 (Goal-AvoidFn (DrsCaseFn DRS24743)) Failure)

(205) (evaluatesOutcome
(PresentationEventFn Sentence-3454360604-73310 IBTGeneration24781)
attack24383 (Goal-MaintainFn protection21295)) Failure)

where the opportunity goal DRS, *DRS24733*, describes the lion harming the ass (shown in Figure 34) and the goal DRS *DRS24743*, which is both successfully achieved and failed to avoid, describes his harming the fox (shown in Figure 33). The relevant fragment of the discourse-level DRS is shown in Figure 38.

(isa see24218 VisualPerception)
(performedBy see24218 lion22287)
(perceivedThings see24218 (DrsCaseFn DRS-3454235720-24775))

DRS-3454235720-24775:

> Universe: ass21039 be24270 trap24296
>
> (isa trap24296 Trapping)
> (objectActedOn trap24296 ass21039)
>
> (isa ass21039 Donkey)

(temporallyIntersects see24218 attack24383)

(isa attack24383 AttackOnObject)
(performedBy attack24383 lion22287)
(objectAttacked attack24383 fox21054)

Figure 38: Partial DRS for the seventh sentence of The Ass, the Fox and the Lion

He then attacked the Ass at his leisure.

The eighth and final sentence presents a goal-directed action on the part of the lion, achieving his goal of harming the ass. The narrative functions inferred are:

(206) (presentsAction
(PresentationEventFn Sentence-3454360606-73632 IBTGeneration247686)
lion22287 attack24540 (Goal-AchieveFn (DrsCaseFn DRS24733)))

(207) (evaluatesOutcome
(PresentationEventFn Sentence-3454360606-73632 IBTGeneration247687)
attack24540 (Goal-AchieveFn (DrsCaseFn DRS24733)) Success)

where the lion's goal DRS, *DRS24733*, describes his harming the ass (shown in Figure 34). The relevant fragment of the discourse-level DRS is shown in Figure 39. Note that the concept of "leisure" in the sentence is not well understood by this interpretation, which focuses on the lion's deliberate action and goal satisfaction.

```
(possessiveRelation lion22287 leisure24626)

(isa attack24540 AttackOnObject)
(performedBy attack24540 lion22287)
(objectAttacked attack24540 ass21039)
```

Figure 39: Partial DRS for the eighth sentence of The Ass, the Fox and the Lion

5.3.3.2 The Dove and the Ant

> An Ant, going to a river to drink, fell in, and was carried along in the stream. A Dove pitied her condition, and threw into the river a small bough, by means of which the Ant gained the shore. The Ant afterward, seeing a man with a fowling-piece aiming at the Dove, stung him in the foot sharply, and made him miss his aim, and so saved the Dove's life.
>
> Little friends may prove great friends.

Figure 40: The Dove and the Ant

An ant went to a river to drink.

She fell into the river and was carried along in the stream.

A dove pitied her condition and threw a small bough into the river.

The ant used the bough to reach the shore.

Afterward, the ant saw a man aiming a gun at the Dove.

The ant stung him in the foot, causing him to miss.

This saved the dove's life.

A little friend may prove a great friend.

Figure 41: QRG-CE version of The Dove and the Ant

Figure 40 contains the original text for the story *The Dove and the Ant* and Figure 41 contains the version rendered in QRG-CE. The complex clauses in the three original sentences are divided into 2-3 sentences each, a few more poetic phrasings and unusual word choices are simplified (e.g. “The Ant afterward…”, “fowling-piece”) and the plural generic in the moral is simplified to a concrete statement.

Figure 42 presents a summary of the event-oriented narrative functions that are interpreted as being used in the narrative (*introducesActor* is omitted). *The Dove and the Ant* is notable for its use of symmetry, which is inferred despite the difference in the presentations of the threats (aware vs. unaware) and the resolution (opportunistic vs. direct outcome).

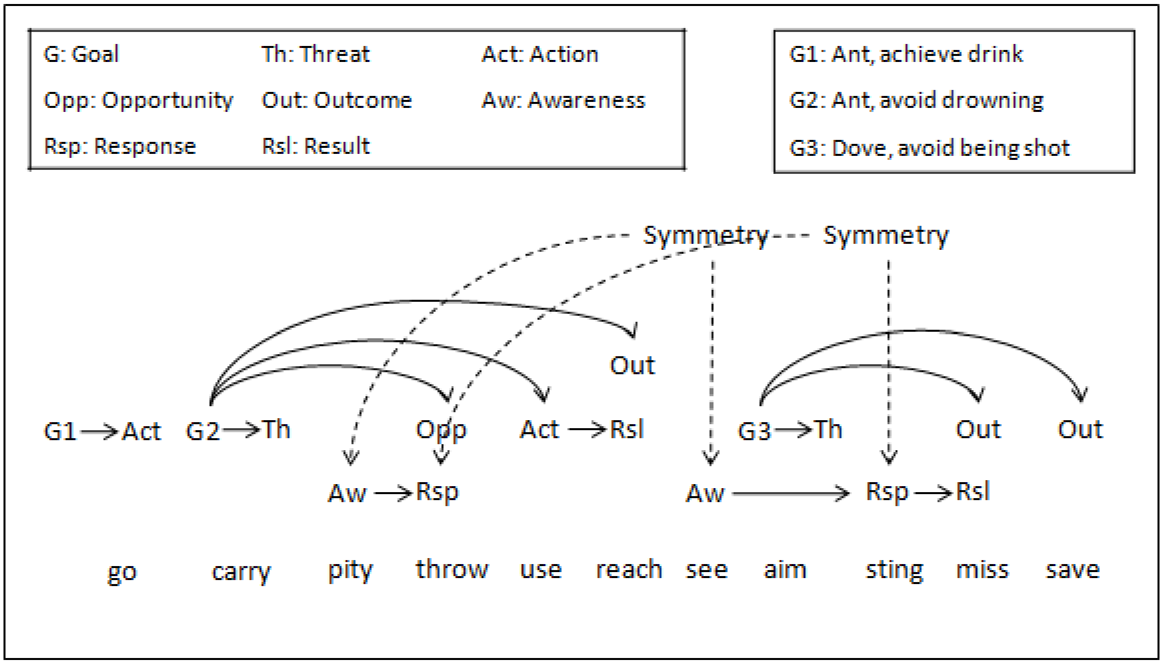

Figure 42: Narrative functions for The Dove and the Ant

5.3.3.3 The Cat and Venus

A CAT fell in love with a handsome young man, and entreated Venus to change her into the form of a woman. Venus consented to her request and transformed her into a beautiful damsel, so that the youth saw her and loved her, and took her home as his bride. While the two were reclining in their chamber, Venus wishing to discover if the Cat in her change of shape had also altered her habits of life, let down a mouse in the middle of the room. The Cat, quite forgetting her present condition, started up from the couch and pursued the mouse, wishing to eat it. Venus was much disappointed and again caused her to return to her former shape.

Nature exceeds nurture.

Figure 43: The Cat and Venus

A cat loved a handsome young man and asked Venus to change her into a woman.

Venus consented to fulfill her request and changed her into a beautiful woman.

This caused the young man to love her and marry her.

Venus wished to discover if the cat, having changed her shape, had changed her habits of life.

So while they were reclining in their chamber, she placed a mouse in the middle of the room.

The Cat, forgetting her present condition, pursued the mouse to eat it.

Venus was very disappointed and returned her to her former shape.

Nature exceeds nurture.

Figure 44: QRG-CE version of The Cat and Venus

Figure 43 contains the original text for the story *The Cat and Venus* and Figure 44 contains the version rendered in QRG-CE. A few more common terms are used (e.g. “asked” vs. “entreated”, “woman” vs. “damsel”), the complex third sentence (“While the two were reclining…”) is divided into two sentences, the semantically complex phase “took her home as his bride” is replaced with the simple “marry her” and the phrase “started up from the couch” is omitted.

Figure 45 presents a summary of the event-oriented narrative functions that are interpreted as being used in the narrative (*introducesActor* is omitted). This story is similar to the other two outcome stories (*The Ass, the Fox and the Lion* and *The Dove and the Ant*) in its heavy use of goal directed actions interwoven with awareness and notable responses. It differs in that there is no explicit threat, only goals being pursued which turn out to conflict in the end.

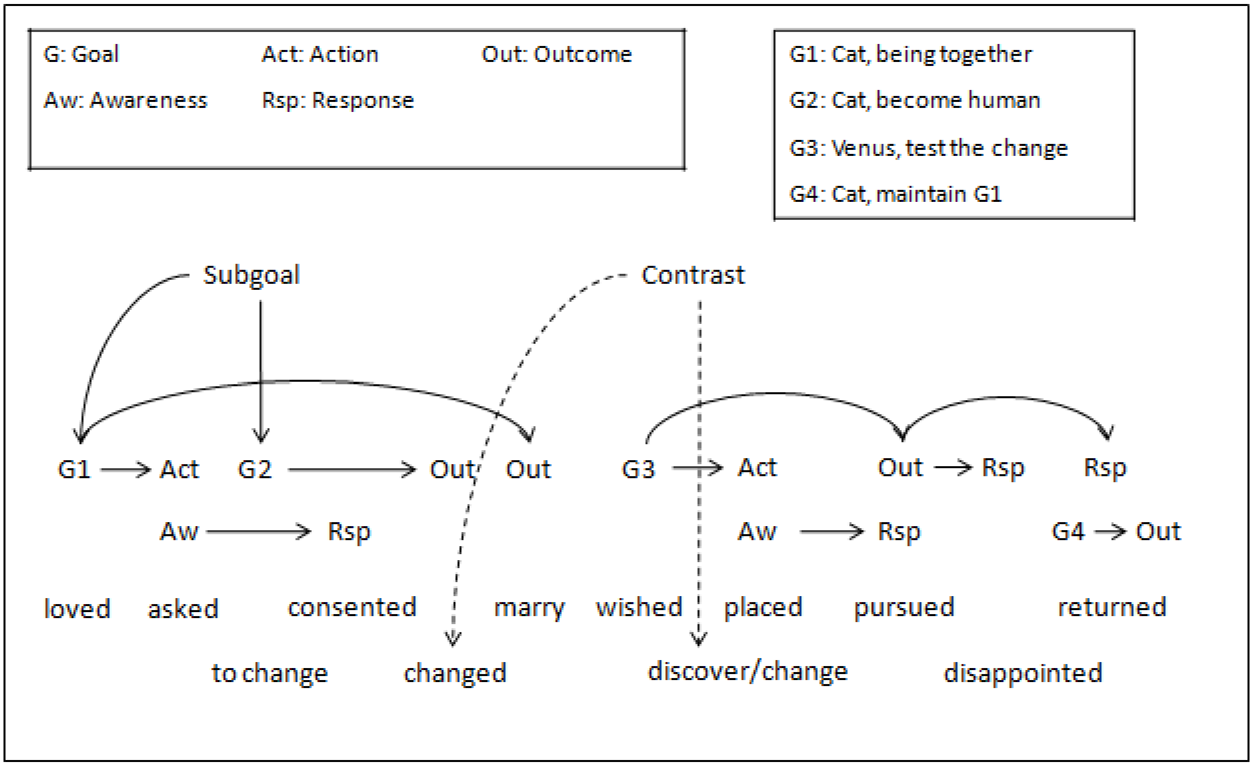

Figure 45: Narrative functions for the Cat and Venus

5.3.3.4 The Dogs and the Fox

SOME DOGS, finding the skin of a lion, began to tear it in pieces with their teeth. A Fox, seeing them, said, "If this lion were alive, you would soon find out that his claws were stronger than your teeth."

It is easy to kick a man that is down.

Figure 46: The Dogs and the Fox

Some dogs, finding the skin of a lion, began to tear it with their teeth.

A fox, seeing them, said, "If this lion was alive, you would soon find that his claws were stronger than your teeth."

It is easy to kick a man that is down.

Figure 47: QRG-CE version of The Dogs and the Fox

Figure 46 contains the original text for the story *The Dogs and the Fox* and Figure 47 contains the version rendered in QRG-CE. This very short story is nearly unchanged except for the omission of the prepositional phrase “in pieces” and the adverbial particle “(find) out”. These alterations avoid adding complexity to the grammar.

Figure 48 presents a summary of the event-oriented narrative functions that are interpreted as being used in the narrative (*introducesActor* is omitted). This narrative does not rely on explicit goals and goal-directed actions to move the plot. Rather, there is a simple pair of observations of awareness followed by responses. The commentary provided by the fox at the end serves to impart meaning to these actions by suggesting a contrasting hypothetical situation where an implicit goal, that of avoiding harm, is violated.

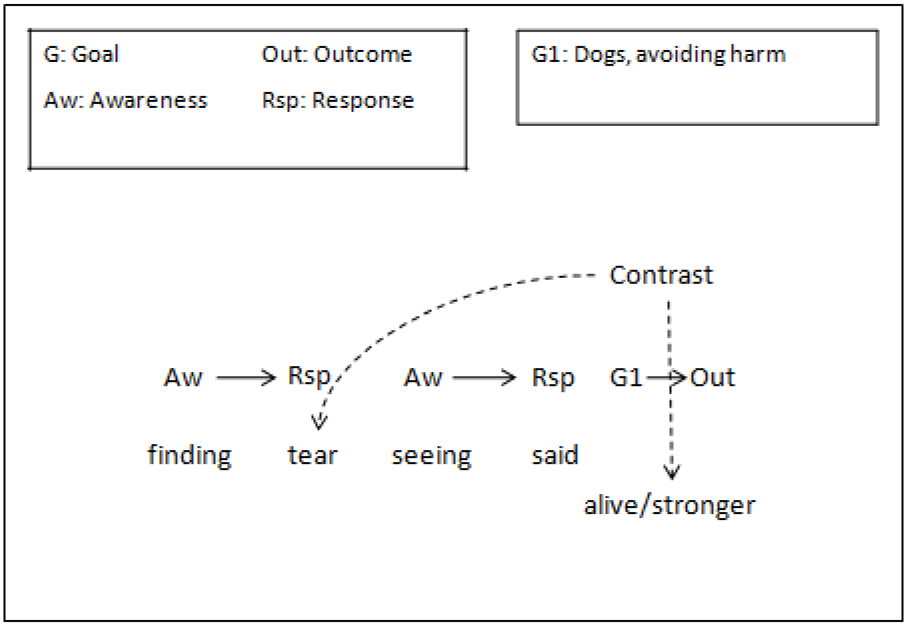

Figure 48: Narrative functions for the Dogs and the Fox

5.3.3.5 The Boy and the Nettles

A BOY was stung by a Nettle. He ran home and told his Mother, saying, "Although it hurts me very much, I only touched it gently." "That was just why it stung you," said his Mother. "The next time you touch a Nettle, grasp it boldly, and it will be soft as silk to your hand, and not in the least hurt you."

Whatever you do, do with all your might.

Figure 49: The Boy and the Nettles

A boy was stung by a Nettle.

He ran to his home and told his mother, "I gently touched it but it very much hurt me."

She said, "That was the reason that it stung you. Whenever you touch a nettle, boldly grasp it and it will not hurt you."

Whatever you do, do with all your might.

Figure 50: QRG-CE version of The Boy and the Nettles

Figure 49 contains the original text for the story *The Boy and the Nettles* and Figure 50 contains the version rendered in QRG-CE. The translation omits the extraneous verb “saying” as well as the poetic embellishment “it will be soft as silk to your hand”. It also reorders the clauses of the boy’s statement to match sequential causality, avoiding the more difficult “although” subordination, and avoids the complex temporal adverbial adjunct “the next time”.

Figure 51 presents a summary of the event-oriented narrative functions that are interpreted as being used in the narrative (*introducesActor* is omitted). Unlike *The Dogs and the Fox,* this narrative does begin with a goal-relevant outcome, that of the boy being harmed. The series of responses and sharing awareness that follows leads to the commentary provided by the mother, which suggests a contrasting approach to what the boy did. The earlier contrast in the story, between the gentleness of the touching and the intensity of the hurting as described by the boy, begins to capture the violation of his expectations. This concept, however, is a more specific relationship than this theory includes.

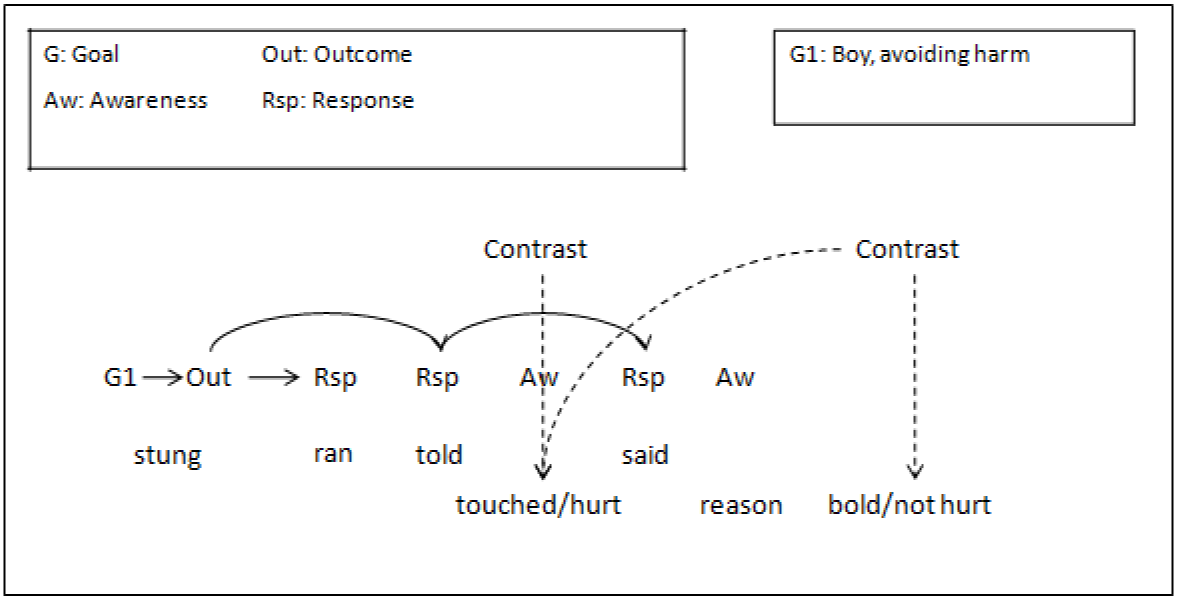

Figure 51: Narrative functions for The Boy and the Nettles

5.3.3.6 The Boys and the Frogs

SOME BOYS, playing near a pond, saw a number of Frogs in the water and began to pelt them with stones. They killed several of them, when one of the Frogs, lifting his head out of the water, cried out: "Pray stop, my boys: what is sport to you, is death to us."

One man's pleasure may be another's pain.

Figure 52: The Boys and the Frogs

Some boys, playing near a pond, saw some frogs and began to pelt them with stones.

They killed several of the frogs.

Then one of the Frogs, lifting his head out of the water, said, "Please stop. Your sport is death to us."

One man's pleasure may be another's pain.

Figure 53: QRG-CE version of The Boys and the Frogs

Figure 52 contains the original text for the story *The Boys and the Frogs* and Figure 53 contains the version rendered in QRG-CE. To reduce parsing complexity, the translation omits the prepositional phrase “in the water” and divides the long second sentence into two sentences to avoid nesting the multi-sentence utterance in a subordinate clause. The poetic “Pray stop, my boys” is replaced with the more straightforward “Please stop.”

Figure 54 presents a summary of the event-oriented narrative functions that are interpreted as being used in the narrative (*introducesActor* is omitted). Like the other two insight-based stories, there is only a single negative outcome, this time placed in the middle of some chains of actions and responses. This story also provides a setting activity, the boys playing, that is not goal-directed like the ant going to get a drink, nor starting a causal chain like the dogs finding the lion’s skin. Rather it provides a background atmosphere that is important to the comparison between sport and death. The frog that provides that commentary at the end is part of the conflict in this story, unlike the previous two, so his response also functions as a goal-directed action. The contrast of concepts in the commentary is explicitly stated.

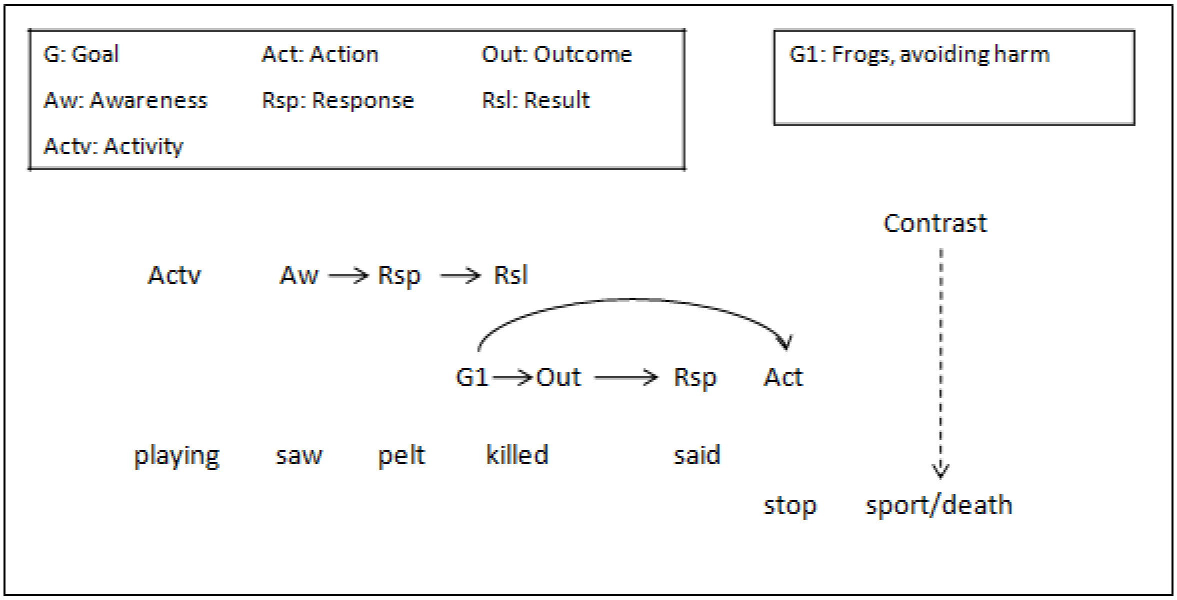

Figure 54: Narrative functions for the Boys and the Frogs

5.3.4 Understanding the morals

The task of moral matching requires an additional theory of how the moral of a fable relates to the story proper. I take the view here that the moral presents a generalization of one interpretation of the story. It takes the sequence of events and opinions given by the story and makes it applicable to a broader context. It is not only boys who should grasp nettles boldly, or only dogs that tear at lions when they are dead. The moral, therefore, makes a statement whose elements are instantiated in the story, and the story as a whole justifies whatever conclusion that statement makes. Interpretation of a moral is proving that it generalizes its story reasonably. The EA NLU task model for this process is implemented in *EAFableMoralQueriesMt* by the single fact:

(208) (queryForInterpretation
0 (narrativeFunction (PresentationEventFn ?sentence-id ?event-id)
NarrativeFunction-Generalize))

The query form is the same as that used in the narrative interpretation task, i.e. trying to prove that the presentation of the sentence serves functional roles in the narrative. However, this query is constrained to only those narrative functions that fall under the category of generalization. This limitation is reasonable in this task because the morals are explicitly given separately from the stories, both in their original form and necessarily for the matching task. Interpretation of the moral sentences proceeds in the same manner as the story proper: the ambiguous, sentence-level representation is generated by the compositional frame semantics and then interpreted and disambiguated in the context of the discourse (the story) using the task defined by expression (208). However, unlike the prior sentences, the resulting interpretation of the moral sentence is not merged into the ongoing discourse. The contents of the moral are not a description of events in the world of the story and thus do not belong in that context. The moral could be situated in the story world as an utterance by an unspecified third party, but nothing would be gained by that here.

This theory of moral generalization has not been expanded beyond the six morals in the moral matching task at hand.

5.3.4.1 Advice morals

Advice morals are identified as imperative statements that suggest that something should either be done or not be done. The DRS representing the action in question, which is easily identified in as the head verb of the imperative statement, provides the qualification of the action. For example, the moral of *The Ass, the Fox and the Lion* is:

(209) Do not trust your enemy.

The disambiguated DRS for this sentence is show in Figure 55.

```
(not (DrsCaseFn DRS-3454063872-26020))

DRS-3454063872-26020:
  Universe: trust25942 your25969 enemy25974

  (isa trust25942 Situation)
  (holdsIn trust25942
   (beliefs (GAP SUBJECT)
    (hasPersonalityTraitToDegree enemy25974 Trustworthy positiveAmountOf)))

  (considersAsEnemy your25969 enemy25974)
  (possessiveRelation your25969 enemy25974)
```

Figure 55: DRS for "Do not trust your enemy."

The negated DRS, *DRS-3454063872-26020*, provides qualification on the situation to be avoided. First, it is characterized by a proposition that holds true about the beliefs of the recipient of the imperative utterance (represented as the syntactic token *(GAP SUBJECT)*). Second, the object of that characterization, represented by the discourse variable *enemy25974* is asserted to have a *considersAsEnemy* relation with the pronoun reference *your25969*, which also resolves to the recipient of the imperative. These qualifications must be exemplified by an event in the story for this moral to apply, and the story must present that event in a negative light for the moral to make sense.

The other advice moral considered in this study comes from *The Boy and the Nettles* which states:

(210) Whatever you do, do with all your might.

This moral states an implication as an imperative, qualifying the latter clause by bindings that satisfy the former. The DRS for this sentence is shown in Figure 56.

(implies-DrsDrs (DrsCaseFn DRS-3454237372-28627) (DrsCaseFn DRS-3454237710-35995))

DRS-3454237372-28627

Universe: do35861 you35857

(isa do35861 PurposefulAction)
(doneBy do35861 you35857)

DRS-3454237710-35995

Universe: do35892

(isa do35892 PurposefulAction)
(doneBy do35892 (GAP SUBJECT))

(coreferent do35892 do35861)

(implies-DrsDrs (DrsCaseFn DRS-3454236165-26186) (DrsCaseFn DRS-3454236165-26187))

DRS-3454236165-26186

Universe: your35930 might35939

(isa might35939 Strength)
(possessiveRelation your35930 might35939)

DRS-3454236165-26187

(instrument-Generic do35892 might35939)

Figure 56: DRS for "Whatever you do, do with all your might."

In the imperative implication, the condition, *DRS-3454237372-28627*, qualifies situations where the consequent holds. Because it is an imperative, this can be understood as an exhortation to make the consequent hold in all situations where the condition holds. That condition involves any *PurposefulAction* done by the pronoun reference *you35857*, which resolves to the recipient of the imperative. The consequent DRS, *DRS-3454237710-35995*, likewise qualifies all *PurposefulActions*, this time done by *(GAP SUBJECT)*, and indicates that this action is a reference to the qualifying action. Thus, for any binding that satisfies the qualifying action, the features asserted in the consequent should be made true. The representation of “with all your might” follows from the syntactic form and, though not entirely intuitive, reasonably captures that everything that satisfies “your might” should be an instrument in the doing. Deeper reasoning could recognize that *Strength* is a conceptual mass noun and infer more domain-specific facts about it. For this moral to apply, a story must exemplify the qualified action and present it in a positive light.

5.3.4.2 Judgment morals

Judgment morals assign a value to a situation or action based on some, often underspecified, scale. This can be accomplished by comparing the relative value of two related elements in the sentence or by relating a single element to a generic quantity. The moral of *The Cat and Venus* states:

(211) Nature exceeds nurture.

This sentence results in the DRS shown in Figure 57.

```
Universe: nature17270 exceed17275 nurture17291

(isa nature17270 IntrinsicForm)
(isa nurture17291 RaisingLivingThings)

(greaterThan-Underspecified nature17270 nurture17291)
```

Figure 57: DRS for "Nature exceeds nurture."

Because work with EA NLU has focused on concrete stories rather than generic rules, concepts tend to be over-instantiated, as in this case. This creates ambiguity as to whether a specific case is being discussed or not. The burden falls on the reasoning task to identify such instances in context, using clues such as the lack of syntactic determiners on the terms "nature" and "nurture" and the lack of role relations for the *nature17270* and *nurture17291* instances in this case. In the same way that reference resolution is handled by contextual reasoning first, then handed to general-purpose heuristic reasoning, EA NLU could implement a set of guidelines for identifying generics. This is out of the scope of this work, but an active area of future work for this approach.

The concepts *IntrinsicForm* and *RaisingLivingThings* are selected to capture the intended meanings of the terms "nature" and "nurture" in this moral. The relationship between them is an underspecified comparison. To generalize a story, this moral requires that elements of the story exemplify the two concepts and that the story provide an instance where the one can be considered greater than the other.

The other type of evaluation moral is seen in *The Dogs and the Fox*, which states:

(212) It is easy to kick a man that is down.

The DRS for this sentence is shown in Figure 58.

```
Universe: it3985 kick4049 be4199 be3990 man4115

(isa kick4049 Kicking)
(performedBy kick4049 (GAP SUBJECT))
(objectActedOn kick4049 man4115)
(isa man4115 UnfortunatePerson)

(degreeOfDifficulty kick4049 Easy)
```

Figure 58: DRS for "It is easy to kick a man that is down."

This moral qualifies a type of action, *Kicking*, performed by (GAP SUBJECT) which does not resolve for the declarative statement and thus is unconstrained. The action is further qualified by the patient role, an instance of an *UnfortunatePerson*. An evaluation is made that this action has a *degreeOfDifficulty* given by the *ScalarInterval Easy*. That is to say, it has a value on a scale of difficulty within the range that denoted by the concept *Easy*. For this moral to apply to a story, it must exemplify the qualified action and provide evidence by example that such an action should be considered easy.

5.3.4.3 Revelation morals

Revelation morals provide surprising wisdom about the state of the world, captured by contrast between the concepts they present. They identify a situation then suggest the possibility that it may lead to a contrasting situation or can be viewed in a contrasting manner. The moral of *The Dove and the Ant* states:

(213) A little friend may prove a great friend.

The DRS for this sentence is shown in Figure 59.

```
Universe: friend20732 AGENT20735

(friends AGENT20735 friend20732)
(sizeParameterOfObject friend20732 (VeryLowToLowAmountFn Size-Generic))

(possible (DrsCaseFn DRS-3454235418-20907))

DRS-3454235418-20907:
    Universe: prove20769 friend20843 AGENT20846

    (friends AGENT20846 friend20843)
    (hasEvaluativeQuantity friend20843 (HighToVeryHighAmountFn Goodness-Generic))

    (isa prove20769 Evidence-Indication)
    (indicated-Prop prove20769 (coreferent friend20732 friend20843))
```

Figure 59: DRS for "A little friend may prove a great friend."

This DRS represents a qualifying situation where there are two entities, *AGENT20735* and *friend20732*, in a *friends* relationship and *friend20732* is qualified as having a particular *sizeParameterOfObject*. Given this situation, the sentence asserts that another situation is *possible* in which there is a similarly qualified “great friend” and evidence is provided that they are one and the same. The two situations show similarity in the friendship relations and contrast in their qualitative features between the low amount and the high amount. The surface level ambiguity of the sentence, omitting whose friend in both noun phrases, is reflected at this level as a reference problem: does *AGENT20846* refer to *AGENT20735*? If the discourse variables in question were explicit in the sentence, this would be a problem of intra-sentential anaphora, which could be resolved. However, EA NLU does not have a general solution to coreference among implicit discourse entities at this time, so this ambiguity persists. For this moral to apply to a story, it must exemplify the qualifying situation of the actor, and then realize the possible future. An event must occur that exemplifies the actor showing him or herself to exemplify the second situation.

The moral of *The Boys and the Frogs* states:

(214) One man's pleasure may be another's pain.

The DRS for this sentence is shown in Figure 60.

```
Universe: pleasure11628 man11605

(isa pleasure11628 Event)
(feelsTowardsEvent man11605 pleasure11628 Pleasure-Feeling mediumToHighAmountOf)
(possessiveRelation man11605 pleasure11628)

(possible (DrsCaseFn DRS-3454234872-11850))

DRS-3454234872-11850:
    Universe: another11722 pain11785

    (isa pain11785 HarmingAnAgent)
    (maleficiary pain11785 another11722)
    (possessiveRelation another11722 pain11785)

    (coreferent pleasure11628 pain11785)
```

Figure 60: DRS for "One man's pleasure may be another's pain."

This moral qualifies an event of any type towards which an actor, *man11605*, feels a positive amount of *Pleasure-Feeling*. Based on this situation, it presents the possibility that there is a *HarmingAnAgent* event occurring and that the two events are in fact the same. This moral generalizes a story which presents an event that exemplifies the qualifying situation, where an actor feels pleasure towards it, that turns out to also exemplify the harm event.

5.3.4.4 Generalizing stories

As discussed above, inferring that a moral generalizes a story requires inferring that elements of the story exemplify elements of the moral. Once those story elements are identified, the form of the moral – advice, evaluation or revelation – can be queried with those story elements replacing the elements they exemplify. In order to capture the intuition that the moral does not come as a complete surprise, but rather confirms and expands a direction the story was already emphasizing, I additionally require that the story elements of interest have been the target of a narrative function of Labovian evaluation, here focusing on contrast and symmetry.

In *The Dove and the Ant*, the story is assumed to be presenting symmetry between the threat/awareness/response patterns that is captured by the inferred narrative functions:

(215) (presentsSymmetry
(PresentationEventFn Sentence-3454063256-18613 IBTGeneration19778)
(aware dove18129 carry17947) (aware ant17670 aim18756))

(216) (presentsSymmetry
(PresentationEventFn Sentence-3454063256-18613 IBTGeneration19780)
(response carry17947 dove18129 throw18205)
(response aim18756 ant17670 sting19379))

Exploring these highlighted events, the act of throwing the branch into the river, *throw18205*, is inferred to exemplify having a little friend as the dove's action befriends the ant by saving her life. The symmetric act of stinging the man in the foot, *sting19379*, is inferred to exemplify being a great friend as the ant's action saved the dove's life. This understanding may lack many subtleties, such as the reciprocal nature of the action, but it is sufficient to identify the application of the moral to the story.

In *The Ass, the Fox and the Lion*, the story is assumed to be presenting a contrast of discrepant awareness when the Fox approaches the Lion to make a deal that he is aware of and the Ass is not. This is captured by the inferred narrative function:

(217) (presentsContrast
(PresentationEventFn Sentence-3454235427-22380 IBTGeneration24740)
(aware fox21054 promise22543) (not (aware ass21039 promise22543)))

shown as expression (192) and repeated here. This contrast provides a potential highlighting of *promise22543*, where the fox promises to capture the ass. This action in the story can be inferred to exemplify the generalization of trusting an enemy that is qualified in the moral. The ultimate outcome for the fox of this action is seen to be negative, which matches with the negative valence of the advice ("do not trust…").

In *The Cat and Venus*, there is an explicit contrast raised by Venus' desire to test between two situations, the actual change of the cat's physical form and the hypothetical change of her routine behaviors. This is captured by the inferred narrative function:

(218) (presentsContrast
(PresentationEventFn Sentence-3454360033-63089 IBTGeneration65500)
(objectOfStateChange change63251 shape63424)
(objectOfStateChange change63630 habit63815))

This contrast provides a potential highlighting of the change events and their target concepts. The judgment moral presented a comparison of the concepts for nature and nurture, which are exemplified by the results of these changes, the cat being a *Cat* versus being a *Person*. Having identified these exemplifications, the interpretation queries for a justification of the *greaterThan-Underspecified* relation between them, and is satisfied by the causal influence of being a cat on the pursuit of the mouse (which is significant due to leading to goal failure) compared to the lack of causal influence of being a person on that event.

In *The Dogs and the Fox*, there is a commentary on the events placed in the mouth of the fox, pointing out an alternative through the explicit conditional "If this lion was alive…". This is captured by a contrast function:

(219) (presentsContrast
(PresentationEventFn Sentence-3454360033-63089 IBTGeneration65500)
(hasExistentialStatus lion52028 Deceased)
(hasExistentialStatus lion52028 Alive))

The hypothetical consequent, where the dogs find the lion equipped to harm them, likewise contrasts with the harm they are doing to the lion in the actual events of the story. This is captured by the contrast function:

(220) (presentsContrast
(PresentationEventFn Sentence-3454360033-63089 IBTGeneration65500)
DRS-3454359746-53090 DRS-3454359776-53105))

The former DRS, *DRS-3454359746-53090*, is inferred to exemplify the generalization in the moral of harming someone when they are in an unfortunate situation. In this moral, unlike the others, the generalization of *Kicking* had to first be generalized to *HarmingAnAgent* before an exemplification could be identified. The judgment that this activity can be attributed the value *Easy* is justified in the story by the contrast between harming another and being harmed oneself.

In *The Boy and the Nettles*, there is an explicit contrast between the gentleness of the boy's touching the nettle and the intensity with which it hurt him. This is captured by the contrast function:

(221) (presentsContrast
(PresentationEventFn Sentence-3454359787-54586 IBTGeneration56397)
(qualityOfAction touch54826 (MediumToHighAmountFn Gentleness))
(qualityOfAction hurt55103 (HighToVeryHighAmountFn Intensity)))

There is also a commentary on the events placed in the mouth of the mother, pointing out an alternative strategy than the boy employed. This is captured by the contrast function:

(222) (presentsContrast
(PresentationEventFn Sentence-3454359787-54586 IBTGeneration56397)
(qualityOfAction touch54826 (MediumToHighAmountFn Gentleness))
(qualityOfAction
grasp56099 (HighToVeryHighAmountFn Courageousness-Feeling)))

The latter contrast highlights the strategy of grasping boldly which exemplifies the moral contention to "…do with all your might." The suggested positive outcome of not being harmed serves to justify the positive valence of the advice.

In *The Boys and the Frogs*, there is a commentary on the events placed in the mouth of one of the frogs, pointing out different perspectives on the same event. This is captured by the contrast function:

(223) (presentsContrast
(PresentationEventFn Sentence-3454359889-59456 IBTGeneration60123)
(isa sport59976 Sport) (isa sport59976 Dying))

The resolution of *sport59976* in the commentary, having a possessive relation to the boys, to the pelting with rocks, *pelt58327*, highlights an event that exemplifies the moral's qualification of *Pleasure*, while also capturing the possible situation of *HarmingAnAgent*.

5.3.5 Matching morals

In this evaluation, six fables were separated into the story proper and the moral sentence. Each fable is processed by EA NLU in the context of the *EANarrativeQueriesMt* microtheory that defines the expectations of narrative functions as a pragmatic reasoning task. This process disambiguates choice sets, including reference resolution, and results in a discourse-level DRS. This DRS contains the facts about the world of the story that are inferred from the text based on the task reasoning. It also results in the set of narrative functions identified for the sentences. After the entire story has been interpreted, each of the six morals is processed. Each moral is appended to the story (as the next sentence in the discourse) and interpreted in the context of the *EAFableMoralQueriesMt* microtheory, described in section 5.3.3.2. If the moral can be assumed through abductive proof to generalize the story, it is selected as an appropriate moral for that story. That moral is then removed from the discourse and each subsequent moral is processed in the same way. In the case of multiple selected morals, the relevance score assigned to each could be used to prioritize them in an order of preference, but that is outside the scope of this evaluation. For all six fables, the system was able to successfully identify the single moral originally provided with the fable ($p < 0.001$).

5.4 Related work

Narrative interpretation was studied by Schank and his colleagues at Yale as a problem of applying world knowledge. They hypothesized that understanding a new story was a matter of invoking previous experiences stored as various types of patterns in memory, then using those patterns to direct subsequent inferences. The FRUMP system (DeJong, 1982) used *scripts* to represent typical scenarios that could then be recognized and summarized in news stories. Wilensky's PAM (Wilensky, 1978) used patterns of causal and intentional behavior to understand actor motivations in utterance pairs. This work demonstrated the effectiveness and necessity of knowledge regarding typical occurrences to make bridging inferences and disambiguate narrative text. It focused coherence within the world of the story, and specifically the coherence of common, everyday experience. This type of coherence is a necessary part of narrative understanding, but it must be subordinate to other factors (i.e. discourse coherence and relevance) otherwise every narrative would be interpreted as being about the most common occurrence that fit its constraints.

Research in story understanding was also quite common in psychology in the 1970s, with theories of *story grammar* being one of the most prominent. Lakoff reformulated Propp's morphology as a grammar using rewrite rules (Lakoff, 1972), and Rumelhart (Rumelhart, 1975) proposed a general grammar aimed at all stories. This was followed by numerous general grammars (Mandler & Johnson, 1977; Stein & Glenn, 1979; Thorndyke, 1977; van Dijk, 1975). These theories were concerned with how human readers store narratives in memory, and how that storage mechanism impacts recall. They were strongly criticized (Black & Wilensky, 1979), defended against those criticisms (Mandler & Johnson, 1980; Rumelhart, 1980) and then criticized again (Garnham, 1983). Black and Wilensky took issue with the ability of a formal grammar to adequately capture the diverse forms found in narrative, and argued that a syntactic characterization of story structure is redundant to semantic knowledge about relationships between story elements. These criticisms were attacked as being fallaciously reached, but Garnham contends that their conclusions are in fact not strong enough. To the first point, he claims that the characterization of story structure as a grammar cannot be supported because no computationally valid account of parsing can be given. To the second point, he claims that there is no justification for imagining a special story-processing mode that would understand or remember stories in a different way than other texts. Rather, that stored knowledge used to understand a story is the same used to understand anything else. I have argued that this knowledge is not sufficient to appreciate the relevant meaning of a story, apart from numerous other coherent interpretations. There are conventions, obligations and expectations in narration that make it more than a mere stand-in for observation.

Later work under Schank at Yale investigated pragmatic expectations in narratives. Lehnert developed a theory of *plot units* as a high-level structure of narrative memory (Lehnert, 1981). While Lehnert focused on the task of automatic summarization, the parallels to research in story grammars for recall tasks are obvious. Lehnert's formulation is simpler than the grammars of Mandler and Johnson or Trabasso… There are three affective states that form the basis of the plot units: *positive event (+)*, *negative event (-)* and *mental state (M),* where M is affectively neutral, and all states are considered with respect to a single character. Four classes of causal links between pairs of states are defined (e.g. motivation, intentionality), resulting in fifteen pair-wise configurations or *primitive plot units*. Lehnert shows by example that compositions of these primitive units can cover a range of familiar plot patterns, and hypothesizes that human readers learn to recognize and use such compositions to divide a narrative into understandable chunks. She shows that these chunks can be identified in human summaries, and used for summary generation, providing evidence that the plot unit structure reflects some aspect of human story understanding. Lehnert argues for a bottom-up rather than top-down approach, criticizing story grammar approaches as having an unrealistic expectation of complete predictive power. The simplicity of Lehnert's formalism, using only the three affective states, is both its strength and its weakness. That a wide range of plot constructs can be captured as causal links between good and bad things happening points to the critical importance of personal outcomes in narrative. Many stories can be summarized as something good or bad happening to someone, with associated reasons why. However, this level of abstraction, categorizing every event as simply positive, negative or neutral, discards a great deal of subtlety and interest. This is particularly true of the categorization of all mental situations as simply neutral.

Dyer took this approach one step further with the proposal of *thematic abstraction units (TAUs)* that represent proverbial knowledge in the form of plan-failure cases (Dyer, 1983). His BORIS implementation used a set of TAUs to do in-depth interpretation of three significantly complex narratives. The TAUs represent familiar lessons that can be used for expectation-driven processing. Dyer showed that if the system is already familiar with the meaning of a story, it can recognize that meaning, and that recognition provides powerful and effective disambiguation and guidance of the interpretation. BORIS integrates previous script and plan work in the form of *memory organization packets (MOPs)* as well as Lehnert's plot units to provide knowledge-rich understanding that can be guided by the application of TAUs. The three narratives read by

BORIS are roughly 20 sentences long, divided into 2 to 6 paragraphs each. They are deliberately similar in an attempt to reuse as much knowledge as possible. BORIS used an expectation-based conceptual parser, *DYPAR*, which was extended from the *McELI* parser (Schank & Riesbeck, 1981). This parser generated *conceptual dependency (CD)* forms (Schank, 1972) based on semantic expectations rather than syntactic constraints. BORIS remains the most in-depth implemented model of computational story understanding, able to answer questions about events, causes, affective impacts and empathetic actions for those three stories. It demonstrated the immense amount of knowledge necessary for detailed story understanding, and there has not been any attempt I am aware of to extend BORIS, or even to apply the significant technologies used to broader scale narrative understanding.

Mooney's work on generalizing novel plans used a similar approach to Dyer, combining schema recognition with plan knowledge to process three shorter stories as well as three stories of three (single clause) sentences each (Mooney, 1988). His evaluation of inferences was more limited than Dyer's, being concerned only with planning knowledge. Following work by Goldman (Goldman, 1991) and Ng (Ng, 1992) on plan recognition in narrative text moved away from schema recognition and used only simple one to four sentence stories. Each sentence presented a single state or event regarding a small set of possible actions surrounding transportation to a supermarket, restaurant, liquor store, park or airport with a suitcase, gun or money and obtaining bread, milk, bourbon or money. Ng processed 50 of these variations, but clearly the focus was firmly on aspects of plan recognition rather than broad inferential relevance. Subsequent research in narrative understanding has tended towards either broad-coverage, shallow inference such as textual entailment and explicit factual QA or deep analysis of isolated phenomena in a handful of example sentences.

Ferguson developed the *MAGI* system for identifying symmetry in structured representations, applying it to imperfect visual line drawings, formal diagrams and narrative (Ferguson, 1994). He used the *structure-mapping engine (SME)* (Falkenhainer et al., 1989) to match a structured description to itself to determine whether two distinct, similar subsections could be found. MAGI was tested on a manually constructed representation of the O. Henry short story *The Gift of the Magi,* and is able to identify the mapping between story events that results in plot symmetry. This algorithm shows robustness across domains, but because SME requires complete representations as input, it was not possible to use it with abduction over ambiguous representations. Using MAGI to robustly identify symmetry on the unambiguous output of EA NLU remains an interesting direction for future work.

5.5 Conclusion

This chapter described a theory of narrative functions and its use with the practical EA NLU approach to guide narrative understanding. The expectation that these functions will appear in narrative is formulated as a query-driven reasoning task, which serves as a pragmatic guide to the EA NLU discourse-level interpretation process. The method used here incrementally disambiguates the discourse as each sentence is added, acting as a heuristic guide for interpretation. This results in a smaller choice space for task reasoning over the entire discourse after the interpretation process is complete.

I have presented the view that narrative is an intentional communication, subject to Gricean discourse pragmatics. In this view, understanding requires inference of story world coherence, discourse coherence and relevance to the reader in narrative understanding. These factors are necessary to explain both the breadth of possible interpretations of a narrative, and the ability to effectively communicate certain intended interpretations. I have argued that current computational accounts of coherence and relevance are insufficient to explain how elements of an ongoing narrative can be judged relevant, and that a heuristic measure of relevance must be introduced. I hypothesize that narrative structure, as understood in the field of narratology, can supply such a heuristic in the form of narrative functions.

The theory of narrative functions presented in this chapter takes inspiration from both narratology and psychological work in memory and recall of stories. I have demonstrated that this theory can be used to guide interpretation of a set of Aesop's fables in EA NLU, and that the resulting interpretations are sufficient to reason about the applicability of the morals of those fables. The use of narrative to illustrate a point is a very general communication capability, and fables including explicit morals are a notable instance of that task. They provide a more well-defined account of the meaning, the intended communicative intent, of each story. That meaning is assumed to have some level of general relevance to readers within this culture. By using EA NLU and narrative functions for this task, I have provided evidence that the approach as a whole is applicable to broader narrative understanding, and that these functions have value as a heuristic for relevance.

I do not claim that this theory of narrative functions is complete. Narrative theory widely recognizes three dimensions of narrative: *action, character* and *setting*. This theory is limited mostly to concerns in the dimension of *action*. Likewise, the domain-specific axioms used to interpret the actions and events in the fables (e.g. falling in a river and the possibility of drowning) are sufficient to cover these stories, but not all the domains that stories involve.

6.0 Conclusion

This dissertation argues for the importance of keeping computational language understanding work connected to the pragmatic use of language. My view is that understanding narrative, for artificial intelligence, should be defined in terms of inferential reasoning ability. This approach requires working with reasonable, well-defined pragmatic concerns. This raises a particular challenge in narrative understanding, where it is difficult to clearly define and scope those concerns. I have argued that cognitive modeling can provide reasoning tasks with clear constraints to fill that need. I have shown this for two particular models, which use text narratives as input and pose non-trivial, real-world reasoning tasks. Cognitive modeling is a novel and useful venue for research in natural language understanding.

In this work I have presented a practical approach to narrative understanding that meets the requirements posed by cognitive models. I have implemented this approach in the EA NLU system, a novel integration of existing natural language understanding theories and resources. EA NLU uses knowledge-rich sub-categorization frames from ResearchCyc for term semantics. The frames express semantic translations in unconstrained CycL, creating a rich and flexible term-level foundation suitable for representing commonsense semantics. The compositional semantics used by EA NLU at the sentence-level supports higher-order compositions of those frames, supporting arbitrarily complex nesting of quantifiers, modal operators, logical connectives and higher-order predicates. This gives EA NLU the full expressive power of CycL for sentence-level representations. The complexity of the compositional process is controlled by the use of controlled grammar, which reduces syntactic ambiguity. Further, it is a context-independent process that admits no general inference over world knowledge. Instead, the composition process chooses least-commitment semantics for some classes of quantifier scope ambiguities, while the remainder of parse, semantic frame, quantifier scope and reference ambiguities are maintained in explicit choice sets. These choice set forms are composed in the same manner as unambiguous forms, and are ultimately exposed to discourse-level processing with context. EA NLU represents these sentence-level compositions as a set of choice sets and a set of dependent discourse representation structures (DRS) adapted from Discourse Representation Theory (Kamp & Reyle, 1993). These structures enable query-driven back-chaining as a discourse interpretation process, allowing pragmatic concerns to be formulated as a set of queries. I have shown that the inputs to the cognitive models considered can be formulated this way, and that the resulting discourse representations are sufficient for the reasoning task they pose. These evaluations used a user intervention model for semi-automatic disambiguation. I further demonstrated that using abductive reasoning for the discourse-level queries automatically performed the same disambiguation for one of those tasks. EA NLU contributes an experimental apparatus, suitable for use in future cognitive modeling experiments with narrative scenarios in natural language. By automating, or even semi-automating, the text encoding process, EA NLU enforces consistent principles of translation and reduces tailorability.

This work also provides evidence that expectations of narrative functions can act as a heuristic for pragmatic judgments of relevance. I have presented a theory of narrative functions and shown that it can be formulated as a pragmatic task in EA NLU to guide interpretation and disambiguation. I have provided evidence that this task-independent guidance leads to representations that are useful for capturing the intended meaning of a set of Aesop's fables. The successful use of this theory in interpretation provides evidence that the EA NLU approach is not limited to only those highly task-specific models used in the earlier evaluations. It also contributes an implementation of a more general pragmatic heuristic for narrative understanding.

6.1 Future work

Allen and colleagues have reported a substantial increase in parsing speed, and a small increase in accuracy, by using output from a shallow, statistical parser (Collins, 1999) to bias probabilities in their deep semantic parser (Swift, Allen, & Gildea, 2004). The current implementation of QRG-CE would not likely see the same benefits, due to the strong constraints of its at-least-one to one design. However, incorporating such biases might allow loosening of the manually constructed constraints in the grammar without an unacceptable loss in performance. This would potentially reduce the effort of adding new rules, by making the grammar more robust to changes, as well as increase its syntactic coverage, making it more user friendly. Of greater interest, however, is the potential of statistical parsing to aid in the interpretation of partial or fragmented parses. It is possible that a shallow parser could be used to "glue together" fragmented parses generated by QRG-CE, providing at least some constraints for semantic role labeling between those fragments. Although such systems use impoverished semantics, the set of semantic roles they use are quite standardized and map well to the syntactic roles in the subcategorization frames. A likely candidate is Bos' BOXER (Bos, 2005) which uses a similar DRT-based representation for quantifier-aware semantic role assignment.

In this work I did not apply semantic heuristics to the disambiguation problem. Cyc provides argument constraints for predicates, which could be used to filter out certain frame choices. However, the nature of language makes such a strict strategy less appealing than softer preferences. These preferences could be implemented and propagated as weights in the abductive reasoning, in the same way Stickel's *weighted abduction* uses manually tuned weights (Stickel, 1989). Automatic assignment of weights to choice sets is a straightforward proposition, allowing any number of knowledge sources to be utilized. Co-occurrence probabilities and term frequencies, both global and with domain-specific, might be applicable through WordNet mappings to Cyc concepts. Knowledge base specific preferences, such as relative ontological positioning, degree of axiomatization and frame redundancy could be applied as well. EA NLU is also able to store manual choices made in the semi-automatic user intervention mode to use as a probabilistic measure of both the quality of frames and their applicability within a domain. These heuristics could also be used to do best-guess disambiguation after the abductive reasoning process has been exhausted. Whether that would help or hurt the usefulness of the representations is an open question.

A significant limitation of the current discourse interpretation model is the lack of accounting for retracting assumptions made in interpreting prior sentences. Within the abductive reasoning process, when a possible new assumption is queried that conflicts with a prior assumption, it is rejected. Thus whatever conclusions might be proven by making that assumption are ruled out. An alternative model could be pursued where those proofs are also seen to completion, but the necessary assumptions attached to that proof cannot be assumed without retracting the prior conflicting ones. Two issues arise. First, it is not clear what effect this would have on the complexity of the abductive proofs. Certainly it would increase the space of possibilities, and it would be competing directly against the gains of incremental interpretation discussed in chapter 5. Second, it is not clear how the system would measure the value of a new assumption, and the proof it justifies, against the value of a conflicting prior assumption and its consequences. This first concern can be empirically tested with minor changes to the current system. The second concern is a general problem in belief revision which has no simple answer. If a retraction was chosen, the most straightforward approach in the current system would be to "roll back" to the sentence with the conflicting assumption, retract it and remove it as a valid choice, then re-interpret the intervening sentences. Alternatively, multiple hypothetical interpretations could be maintained simultaneously with an ATMS (de Kleer, 1986), eliminating the need for general belief revision. While ATMS algorithms suffers from combinatorial explosion, a *Hybrid Truth Maintenance System (HTMS)* (de Kleer, 1994) restricts label propagation to a set of *focus environments*. Because EA NLU can easily identify conflicting sets of assumptions in the abductive proof, it is likely that focus environments could be defined during the interpretation process. The LTMS working memory currently used could be replaced by an HTMS for testing. The empirical question of how many environments would be generated, and in particular how well it would scale with story length and the number of choice sets, could then be tested.

Work is already underway exploring the use of public lexical-semantic resources to accelerate or automate adding sub-categorization frames to the knowledge base. The first difficulty is creating the mappings from plain English terms to concepts in Cyc. WordNet (Fellbaum, 1998) provides semantic similarity information in the form of *synsets*, clusters of similar word senses, that suggest common mappings for terms. This allows an existing mapping to be replicated for other terms, but would clearly be promiscuous because each term has multiple senses. WordNet provides example sentences for each word sense, which might be compared against concept names and document strings in Cyc to try to constrain this process. Alternatively, these automatically generated mappings might be used in a semi-automatic tool to reduce the amount of knowledge base searching required of the user. VerbNet (CITE) maps verbs into a hierarchy of classes that specific applicable thematic roles and selectional constraints on those roles, in addition to syntactic constraints. These classes can be translated into sub-categorization frames by mapping the thematic roles to high-level role relations. To the extent that the selectional restrictions can be mapped to Cyc concepts, they could be compared with argument constraints in Cyc to move from those high-level role relations down the predicate ontology to more specific role relations. The challenge of mapping from a verb class to a Cyc concept is the same as mapping from terms, but it is likely that several members of any given class do have defined frames. It is possible that an appropriate concept for a class could be found by searching the most specific common generalizations of concepts in those frames.

Ongoing projects are using EA NLU with cognitive models of conceptual change, tutoring dialogues about commonsense knowledge and multimodal learning from science textbooks. EA NLU is able to support such diverse input due to the task-independent flexibility of the subcategorization frames, compositional frame semantics and query-driven interpretation. The use of EA NLU across tasks and domains provides important pressure on QRG-CE to expand coverage while remaining task and domain independent.

The theory of narrative functions requires increased coverage to act as a general-purpose narrative pragmatics for interpretation. In particular, the narrative dimensions of character and setting received little attention in the current model. The functions of setting situate the action in space and time. As a measure of relevance, statements that introduce set pieces and props are of particular interest because they raise the expectation that those entities will be used in a notable way; a river might be fallen into or something might be purchased in a town. Possible functions are:

(224) (introducesSetting ?setting)

(225) (introducesProp ?prop)

(226) (presentsUse ?setting-or-prop ?usage)

where the variable *?usage* is a relation between the setting or prop and an event. Typical relations might include *instrument*, *enables* and *into*. Characterization is the function of indicating that an actor possesses a certain internal trait. This characterization can be diegetic (explicitly told by the narrator) or mimetic (indicated within the story itself). Figure 61 shows the typical categories of mimetic characterization such as proposed by Tomashevsky (Tomashevsky, 1925/1965).

Explicit (direct)
- Expository self-description
- Expository description of another

Implicit (indirect)
- External appearance
- Non-verbal behavior
- Verbal behavior
    - Content of utterance

Figure 61: categories of mimetic characterization

A function for characterization might be:

(227) (characterize ?category ?subcat ?actor ?trait)

where the *?category* and *?subcat* variables are bound to constants indicating the means of characterization and *?trait* is a character trait concept. Character traits are more-or-less persistent qualities such as being patient or compassionate rather than temporary emotions such as happiness or frustration. Work with other forms of narratives beyond fables, such as personal recounting, will provide evidence as to the generality of this theory of narrative functions as a heuristic for relevance and help to refine the theory moving forward.

6.2 Final thoughts

Narrative understanding with natural language is a very hard problem that the research community will be working on for some time to come. This work contributes an implemented model of deep language understanding with a clearly defined interface to pragmatic concerns. Also, it is grounded in a large-scale, widely available knowledge base and ontology, making it applicable to a wide range of pragmatic reasoning tasks. I believe that this approach is important to future work in narrative understanding, and language understanding in general. Connecting independently motivated reasoning tasks to language understanding provides important constraints and goals for both endeavors.

7.0 References

Allen, J., Chambers, N., Ferguson, G., Galescu, L., Jung, H., Swift, M., et al. (2007). *PLOW: A collaborative task learning agent*. Vancouver, BC.

Allen, J. F. (1994). *Natural Language Understanding* (2nd ed.). Redwood City, CA: Benjamin/Cummings.

Allen, J. F., Byron, D. K., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. (2001). Towards Conversational Human-Computer Interaction. *AI Magazine, 22*, 27-37.

Allen, J. F., & Core, M. (1997). *DAMSL: Dialog act markup in several layers*. Unpublished manuscript.

Allen, J. F., & Ferguson, G. (1994). Actions and events in interval temporal logic. *Journal of Logic and Computation*.

Asher, N., & Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

Austin, J. (1975). *How to Do Things With Words*. Cambridge, MA: Harvard University Press.

Banko, M., Cafarella, M., Soderland, S., Broadhead, M., & Etzioni, O. (2007). *Open information extraction from the web*. Hyderabad, India.

Baron, J., & Spranca, M. (1997). Protected Values. *Organizational Behavior and Human Decision Processes, 70*, 1-16.

Barthes, R. (1977). Introduction to the structural analysis of narratives (S. Heath, Trans.). In *Image, music, text*. New York: Hill and Wang.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Bawakid, A., & Oussalah, M. (2008). *A semantic summarization system: University of Birmingham at TAC 2008*.

Bertens, H. (2008). *Literary Theory: the basics*. Taylor & Francis.

Black, J. B., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science, 3*, 213-230.

Boolos, G. (1984). To be is to be the value of a variable (or to be some values of some variables). *Journal of Philosophy, 81*, 430-449.

Bos, J. (2005). *Towards wide-coverage semantic interpretation*. Tilburg, The Netherlands.

Boutilier, C., & Becher, V. (1995). Abduction as Belief Revision. *Artificial Intelligence*, *77*(1), 43-94.

Bratman, M. (1990). What is intention? In P. Cohen, J. Morgan & M. Pollack (Eds.), *Intentions in Communication*. Boston, MA: MIT Press.

Bylander, T., Allemang, D., Tanner, M. C., & Josephson, J. R. (1991). The computational complexity of abduction. *Artificial Intelligence*, *49,* 25-60.

Charniak, E. (1972). *Toward a model of children's story comprehension*. Unpublished manuscript.

Charniak, E. (1977). A framed PAINTING: the representation of a common sense knowledge fragment. *Cognitive Science*.

Charniak, E., & Goldman, R. (1988). *A logic for semantic interpretation*. Buffalo, NY.

Charniak, E., & Goldman, R. (1989). *A semantics for probabilistic quantifier-free first-order languages, with particular application to story understanding*. Detroit, MI.

Chatman, S. (1978). *Story and Discourse: Narrative Structure in Fiction and Film*. Ithaca: Cornell University Press.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse Production and Comprehension*. Norwood, NJ: Ablex.

Clark, P., & Harrison, P. (2008). *Boeing's NLP System and the Challenges of Semantic Representation*. Venice, Italy.

Clark, P., Harrison, P., Jenkins, T., Thompson, J., & Wojcik, R. (2005). *Acquiring and Using World Knowledge using a Restricted Subset of English*.

Collins, M. (1999). *Head-driven statistical models for natural language parsing.* University of Pennsylvania.

Cox, P. T., & Pietrzykowski, T. (1986). *Causes for events: their computation and applications*. Oxford, England.

Cullingford, R. E. (1978). *Script application: Computer understanding of newspaper stories.* Yale University.

Curits, J., Witbrock, M., Baxter, D., Cabral, J., Wagner, P., & al., e. (2006). *The TextLearner System: Reading Learning Comprehension Final Report*. Unpublished manuscript.

Curran, J. R., Clark, S., & Bos, J. (2007). *Linguistically Motivated Large-Scale NLP with C&C and Boxer*.

Curtis, J., Baxter, D., Wagner, P., Cabral, J., Schneider, D., Witbrock, M., et al. (2009). Methods of Rule Acquisition in the TextLearner System. *Proceedings of the 2009 AAAI Spring Symposium on Learning by Reading and Learning to Read*. Palo Alto, CA: AAAI Press.

Dang, H. T., Kelly, D., & Lin, J. (2007). *Overview of the TREC 2007 Question Answering Track*. (Vol. SP 500-274): NIST.

Dauge, P. (1993). Numeric Reasoning with Relative Orders of Magnitude. *Proceedings of the Seventh International Workshop on Qualitative Reasoning*.

Davidson, D. (2001). The logical form of action sentences. *Essays on Actions and Events*, *1*(9), 105-149.

de Kleer, J. (1986). An assumption-based TMS. *Artificial Intelligence*, *28*, 127-162.

de Kleer, J. (1994). *A Hybrid-Truth Maintenance System.* PARC

Dehghani, M., Gentner, D., Forbus, K., Ekhtiari, H., & Sachdeva, S. (2009). *Analogy and Moral Decision Making*.

Dehghani, M., Tomai, E., Forbus, K., & Klenk, M. (2008). An Integrated Reasoning Approach to Moral Decision-Making. *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*. Chicago, IL.

DeJong, G. (1982). An overview of the frump system. In W. Lehnert & M. Ringle (Eds.), *Strategies for Natural Language Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

deKleer, J. (1986). An assumption-based TMS. *Artificial Intelligence*, *28*, 127-162.

Downey, D., Etzioni, O., & Soderland, S. (2005). *A Probabilistic Model of Redundancy in Information Extraction*. Hyderabad, India.

Dyer, M. (1983). *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension*. Cambridge, MA: MIT Press.

Eshghi, K. (1992). *Tractable abduction*. Unpublished manuscript.

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). The structure mapping engine: algorithm and example. *Artificial Intelligence, 41*, 1-63.

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Ferguson, R. (1994). *MAGI: Analogy-based encoding using regularity and symmetry*.

Fillmore, C. (2006). Frame Semantics. In D. Geeraerts (Ed.), *Cognitive Linguistics: Basic Readings*. Berlin, New York: Mouton de Gruyter.

Forbus, K. (1984). Qualitative process theory. *Artificial Intelligence, 24*, 85-168.

Friedman, S., & Forbus, K. (2008). *Learning Causal Models via Progressive Alignment & Qualitative Modeling: A Simulation*. Washington, D.C.

Garnham, A. (1983). What's wrong with story grammars. *Cognition*, *15*(1-3), 145-154.

Goldman, R. P. (1991). *A Probabilistic Approach to Language Understanding.* Brown University.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3). New York, NY: Academic Press.

Grishman, R., Macleod, C., & Wolff, S. (1993). *The COMLEX Syntax Project*. Ft. Belvoir: Defense Technical Information Center.

Heider, F. (1983). *The psychology of interpersonal relations*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hobbs, J. (1985). *On the coherence and structure of discourse.* Center for the Study of Language and Information.

Hobbs, J. (1986). Overview of the TACITUS project. *Proceedings of Strategic Computing - Natural Language Workshop*. Marina del Rey, CA.

Hobbs, J., Stickel, M., Appelt, D., & Martin, P. (1990). Interpretation as Abduction. *Artificial Intelligence*, *63*, 69-142.

Hovy, E. H. (1988). *Planning coherent multisentential text*. Buffalo, NY.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Prentice Hall series in artificial intelligence. Upper Saddle River, N.J.: Pearson Prentice Hall.

Kamp, H., & Reyle, U. (1993). *From discourse to logic : introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Studies in linguistics and philosophy ; v. 42. Dordrecht ; Boston: Kluwer Academic.

Kaplan, R., Riezler, S., King, T. H., Maxwell, J. T. I., Vasserman, A., & Crouch, R. (2004). *Speed and accuracy in shallow and deep stochastic parsing*. Boston, MA.

Kim, S. (1999). Causal briding inference: A cause of story interestingness. *British Journal of Psychology*, *90*, 57-71.

Kinsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Knobe, J. (2003). Intentional Action and Side-Effects in Ordinary Language. *Analysis*, *63*, 190-193.

Kuehne, S. (2004). *Understanding natural language descriptions of physical phenomena*. Northwestern University, Evanston, IL.

Labov, W., & Waletzky, J. (1966). Narrative Analysis: Oral Versions of Personal Experience. In J. Helms (Ed.), *Essays on the Verbal and Visual Arts* (pp. 12-44). Seattle: University of Washington Press.

Lakoff, G. P. (1972). Structural complexity in fairy tales. *The study of man*, *1*, 128-190.

Lehnert, W. (1981). Plot Units and Narrative Structure. *Cognitive Science*.

Lehnert, W. (1994). Cognition, computers and car bombs: How Yale prepared me for the 1990s. In R. Schank & E. Langer (Eds.), *Beliefs, reasoning and decision making*. Hillsdale, NJ: Erlbaum.

Lenat, D., & Gupta, R. V. (1990). *Building Large Knowledge Based Systems*. Reading, MA: Addison Wesley.

Lewis, J. E. (1996). *The English Fable: Aesop and Literary Culture, 1651-1740* (H. Erskine-Hill & J. Richetti, Ed.). Cambridge Studies in Eighteenth-Century English Literature and Thought. Cambride University Press.

Lim, C. S., & Baron, J. (1997). *Protected values in Malaysia, Singapore and the United States*. Unpublished manuscript.

Macleod, C., Grishman, R., & Meyers, A. (1998). *COMLEX: Syntax Reference Manual, Version 3.0*. Pennsylvania, PA: Linguistic Data Consortium, University of Pennsylvania.

Mandler, J. M., & Johnson, N. S. (1977). Rememberance of things parsed. *Cognitive Psychology*, *9*, 111-151.

Mandler, J. M., & Johnson, N. S. (1980). On throwing out the baby with the bathwater: A reply to Black and Wilensky's evaluation of story grammars. *Cognitive Science*, *4*(3), 305-312.

Mann, W., & Johnson, M. (1986). Relational propositions in discourse. *Discourse Processes*, *9*(1), 57-90.

Mao, W. (2006). *Modeling Social Causality and Social Judgment in Multi-Agent Interactions.* University of Southern California, Los Angeles, California.

Matsuzaki, T., Miyao, Y., & Tsujii, J. (2007). *Efficient HPSG parsing with supertagging and CFG-filtering*. Hyderabad, India.

Mooney, R. J. (1988). *A general explanation-based learning mechanism and its application to narrative understanding.* University of Illinois at Urbana-Champaign.

Mross, E., & Roberts, J. (1992). *The construction integration model: A program and a manual.* Boulder, CO: Institute for Cognitive Science, University of Colorado.

Nair, R. B. (2003). *Narrative gravity: conversation, cognition, culture*. Routledge.

Ng, H. T. (1992). *A general abductive system with application to plan recognition and diagnosis.* University of Texas, Austin.

Ng, H. T., & Mooney, R. J. (1990). *The role of coherence in constructing and evaluating abductive explanations*.

Norvig, P. (1987). *Inference in text understanding*. Seattle, WA.

Norvig, P., & Wilensky, R. (1990). *A critical evaluation of commensurable abduction models for semantic interpretation*. Helsinki, Finland.

Nyberg, E., Mitamura, T., Baker, K., Svoboda, D., Peterson, B., & Williams, J. (2002). *Deriving Semantic Knowlege from Descriptive Texts Using an MT System.* Tiburon, CA.

Paul, G. (2000). AI approaches to abduction. In D. M. Gabbay & R. Kruse (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems* (Vol. 4). Netherlands: Kluwer Academic Publishers.

Pearl, J. (1987). Distributed revision of composite beliefs. *Artificial Intelligence*, *33*(2), 173-215.

Peirce, C. S. (1955). *Philosophical writings of Peirce*. New York,: Dover Publications.

Pople, H. E. (1973). *On the Mechanization of Abductive Logic*. Stanford, CA.

Pradhan, S., Ward, W., Hacioglu, K., Martin, J., & Jurafsky, D. (2004). *Shallow semantic parsing using support vector machines*.

Prince, G. (1983). Narrative pragmatics, message and point. *Poetics*, *12*, 527-536.

Propp, V. (1971). *Morphology of the Folktale*. Austin: University of Texas Press.

Ramachandran, D., Reagan, P., & Goolsbey, K. (2005). *First-Orderizing ResearchCyc: Expressivity and Efficiency in a Common-Sense Ontology*. Pittsburgh, PA.

Reggia, J. A., Nau, D. S., & Wang, P. (1983). Diagnostic expert systems based on a set covering model. *International Journal of Man-Machine Studies*, *19*(5), 437-460.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Free Press.

Riezler, S., King, T. H., Crouch, R., & Zaenen, A. (2003). *Statistical sentence condensation using ambiguity packing and stochastic disambiguation methods for Lexical-Functional Grammar*. Edmonton, Canada.

Ritov, I., & Baron, J. (1999). Protected Values and Omission Bias. *Organizational Behavior and Human Decision Processes*, *79*(2), 79-94

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: studies in cognitive science*. New York: Academic Press.

Rumelhart, D. E. (1980). On evaluating story grammars. *Cognitive Science*, *4*(3), 313-316.

Ryan, M. L. (1992). *Possible Worlds, Artificial Intelligence, and Narrative Theory*. Bloomington: Indiana University Press.

Schank, R. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, *3*(4), 552-631.

Schank, R., & Ableson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum.

Schank, R., & Riesbeck, C. (1981). *Inside computer understanding: five programs plus miniatures*. Hillsdale, NJ: Lawrence Erlbaum.

Schlaefer, N., Ko, J., Betteridge, J., Sautter, G., Pathak, M., & Nyberg, E. (2007). *Semantic extensions of the Ephyra QA system for TREC 2007*. NIST.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes*, *33*(2), 105-132.

Schubert, L. K., & Hwang, C. H. (2000). Episodic Logic Meets Little Red Riding Hood: A Comprehensive, Natural Representation for Language Understanding. In L. Iwanska & S. C. Shapiro (Eds.), *Natural Langauge Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (pp. 111-147). Menlo Park, CA and Cambridge, MA: MIT/AAAI Press.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, UK: Cambridge University Press.

Searle, J. R. (1975). Indirect Speech Acts. In Cole & Morgan (Eds.), *Syntax and Semantics* (Vol. 3). New York: Academic Press.

Searle, J. R. (1979). Literal meaning. In J. R. Searle (Ed.), *Expression and Meaning*. Cambridge, MA: Cambridge University Press.

Selman, B., & Levesque, H. J. (1990). *Abductive and default reasoning: A computational core*.

Shaver, K. G. (1985). *The Attribution Theory of Blame: Causality, Responsibility and Blameworthiness*. Springer-Verlag.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, *31*, 507-524.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing*. Norwood, NJ: Ablex.

Stickel, M. (1989). Rationale and methods for abductive reasoning in natural-language interpretation. *Lecture Notes in Artificial Intelligence, 459*, 233-252.

Suh, S. Y., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32*, 279-300.

Swift, M., Allen, J., & Gildea, D. (2004). *Skeletons in the parser: using a shallow parser to improve deep parsing*. Geneva, Switzerland.

Tetlock, P. E. (2000). Cognitive biases and organizational correctives: Do both disease and cure depend on the ideological behavior? *Administrative Science Quarterly, 45*(2), 293-326.

Thorndyke, P. W. (1977). Cognitive structures in comprehesion and memory of narrative discourse. *Cognitive Psychology, 9*, 77-110.

Tomai, E., & Forbus, K. D. (2008). Using Qualitative Reasoning for the Attribution of Moral Responsibility. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*. Washington, DC.

Tomashevsky, B. (1925/1965). Thematics. In L. T. Lemon & M. J. Reis (Eds.), *Russian formalist criticism: Four essays*: Lincoln.

Trabasso, T., Secco, T., & van den Broek, P. W. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83-111). Hillsdale, NJ: Lawrence Erlbaum Associates.

van den Broek, P., Lorch, R. F. J., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory and Cognition, 29*(8), 1081-1087.

van Dijk, T. A. (1975). *Recalling and summarizing complex discourse*. Unpublished manuscript, Amsterdam.

Waldmann, M. R., & Dieterich, J. (2007). Throwing a Bomb on a Person Versus Throwing a Person on a Bomb: Intervention Myopia in Moral Intuitions. *Psychological Science, 18*(3), 247-253.

Weiner, B. (1995). *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. Guilford Press.

Weiner, B. (2001). Responsibility for Social Trangressions: An Attributional Analysis. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. Boston, MA: MIT Press.

Wilensky, R. (1978). *Understanding Goal-Based Stories.* Yale University.

Wilensky, R. (1983). *Planning and Understanding: A Computational Approach to Human Reasoning*. Reading, MA: Addison-Wesley.

Wilensky, R., Chin, D. N., Luria, M., Martin, J., Mayfield, J., & Wu, D. (2000). The Berkeley UNIX Consultant Project. *Artificial Intelligence Review*, *14*(1-2), 43-88.

Wilson, D., & Matsui, T. (2000). Recent approaches to bridging: Truth, coherence, relevance. In J. de Bustos Tovar, P. Charaudeau, J. Alconchel, I. Recuero & C. L. Alonso (Eds.), *Lengua, Discurso, Texto* (Vol. 1, pp. 103-132). Madrid: Visor Libros.

Wilson, D., & Sperber, D. (2004). Relevance Theory. In *Handbook of Pragmatics*. Oxford: Blackwell.

8.0 Appendices

8.1 Appendix A: QRG-CE

(Grammar rules specific to QP patterns and question answering omitted)

```
;;;; -*-  Mode: LISP; Syntax: Common-Lisp; Base: 10                          -*-
;;;; ---------------------------------------------------------------------------
;;;; File name: qrgce-grammar2-25.dat
;;;;    System: EA
;;;;    Author: Emmett Tomai
;;;;   Created: June 12, 2007
;;;;   Purpose:
;;;; ---------------------------------------------------------------------------

(headfeatures
 (slp vform perfect progressive agr)
 (rel vform agr)
 (vp subcat inf vsay np-vsay vmotion parenthetical)
 (cnp nunit countable ntime1 aggregate ncollective)
 (np ntime1 nunit)
 (advp modif)
 (aux agr))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Experimental Section
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; compound noun
((cnp (var ?varcnp2) (agr ?a2)
      (sem (d::and ?semcnp1 ?semcnp2
                   (d::compoundNoun ?varcnp1 ?varcnp2))))
 -cnp->cnp-cnp- 0.8
 (cnp (var ?varcnp1) (agr ?a1) (sem ?semcnp1) (gerund -) (prep-lex -) (modif-adj -))
 (head (cnp (var ?varcnp2) (agr ?a2) (sem ?semcnp2) (gerund -) (prep-lex -) (modif-adj -))))

;;; descriptive pname proper noun "Bosnian buses"
((cnp (var ?varcnp2) (agr ?a2)
      (sem (d::and ?semcnp2 (d::possessiveRelation ?semn ?varcnp2))))
 -cnp->pname-cnp- 0.8
 (pname (agr ?a) (var ?varn) (lex ?lexn) (orth ?orthn) (sem ?semn))
 (head (cnp (var ?varcnp2) (agr ?a2) (sem ?semcnp2) (gerund -) (prep-lex -) (modif-adj -))))

;;; unknown as proper name
((np (agr (? a 1s 2s 3s 1p 2p 3p)) (var ?semu) (assumed-pname ?semu)
     (sem (d::thereExists ?semu)))
 -np->unknown-as-pname- 0.9
 (head (unknown (var ?varu) (sem ?semu))))

((np (agr (? a 1s 2s 3s 1p 2p 3p)) (var ?semu) (assumed-pname ?semu)
     (sem (d::thereExists ?semu)))
 -np->the-unknown-as-pname- 0.9
 (det (lex the))
 (head (unknown (var ?varu) (sem ?semu))))

;;; noun aside "Hossein, a 13 year old boy, ..."
((np (var ?varnp) (agr ?a)
     (sem (d::and ?semnp ?semaside
                  (d::denotes ?varnp ?varaside))))
 -np->noun-comma-np-aside-comma- 0.9
 (head (np (var ?varnp) (sem ?semnp) (agr ?a)))
```

```
 (punc (lex punc-comma))
 (np (var ?varaside) (sem ?semaside) (agr ?a))
 (punc (lex punc-comma)))

;;; noun-be-that-s (clausal substitution)
((cnp (var ?varn) (agr ?a)
      (:NOUN ?varn)
      (:CLAUSE ?semslp)
      (sem ?semn))
 -cnp->noun-that-s-
 (head (noun (var ?varn) (sem ?semn) (agr ?a) (subcat noun-be-that-s)))
 (sconj (lex that))
 (slp (var ?varslp) (sem ?semslp)))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Common Noun Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; noun as common noun phrase
((cnp (var ?varn) (agr ?a) (ntime1 ?nt1) (ntime2 ?nt2)
      (:NOUN ?varn)
      (sem ?semn))
 -cnp->n-
 (head (noun (sem ?semn) (agr ?a) (var ?varn) (ntime1 ?nt1) (ntime2 ?nt2))))

;; "red car", "hot day"
((cnp (var ?varcnp) (agr ?a) (modif-adj +) (globally-quantified ?gquant)
      (:NOUN ?varcnp)
      (:REPLACE (d::ConceptForInstanceFn ?semcnp))
      (sem (d::and ?semcnp ?semadjp)))
 -cnp->adjp-cnp- 0.9
 (adjp (sem ?semadjp) (var ?varadjp) (global-quantifier ?gquant))
 (head (cnp (var ?varcnp) (agr ?a) (sem ?semcnp) (gerund -))))

;;; "5 foot stick"
((cnp (var ?varn) (agr ?a)
      (:NOUN ?varn)
      (sem (d::and (d::measure-Underspecified ?varn ?semmp) ?semn)))
 -cnp->mp-n-
 (mp (sem ?semmp) (var ?varmp))
 (head (noun (var ?varn) (agr ?a) (sem ?semn))))

;;; "5 liters of water"
;; hasAmount is for QP stuff, measure might be more general
((np (var ?varpp) (agr ?a) (generic +)
     (:NOUN :NO-NOUN)
     (sem (d::thereExists ?varpp
           (d::and ?sempp
                   (d::measure ?varpp ?semmp)
                   (d::hasAmount :SUBJECT ?semmp)))))
 -np->mp-pp-
 (mp (sem ?semmp) (var ?varmp) (agr ?a))
 (head (pp (prep-lex of) (var ?varpp) (sem ?sempp))))

;; "bunch of grapes", "temperature of the brick", "ball with stripes"
((cnp (var ?varcnp) (agr ?a)
      (:NOUN ?varcnp)
      (:POSSESSOR ?varpp)
      (:OBLIQUE-OBJECT ?varpp)
      (prep-lex ?!plex)
      (sem (d::and ?semcnp ?sempp)))
 -cnp->cnp-pp-
 (head (cnp (var ?varcnp) (sem ?semcnp) (agr ?a) (gerund -) (nunit -) (modif-adj -)))
 (pp (var ?varpp) (sem ?sempp) (prep-lex (? !plex into to during at for by out-of))))

((cnp (var ?varcnp) (agr ?a)
      (:NOUN ?varcnp)
```

```
      (:POSSESSOR ?varpp)
      (:OBLIQUE-OBJECT ?varpp)
      (prep-lex ?plex)
      (sem (d::and ?semcnp ?sempp)))
 -cnp->cnp-pp-unlikely- 0.9
 (head (cnp (var ?varcnp) (sem ?semcnp) (agr ?a) (gerund -) (nunit -) (modif-adj -)))
 (pp (var ?varpp) (sem ?sempp) (prep-lex (? plex into to during at for by out-of))))

;;; not sure how to constrain this one, "opportunity" doesn't have
;; special comlex tags
((cnp (var ?varn) (agr ?a)
      (:SUBJECT ?varn)
      (:NOUN ?varn)
      (sem (d::and ?semn ?seminf)))
 -cnp->n-npinf-
 (head (noun (var ?varn) (sem ?semn) (agr ?a)(lex opportunity)))
 (np (var ?varinf) (sem ?seminf) (inf +) (np-inf -)))

;;;;;;;;;;;;
;; complex clause cnp

;; "man sitting on the floor"
((cnp (var ?varcnp) (agr ?a)
      (:SUBJECT ?varcnp)
      (:OBJECT :OBLIQUE-OBJECT)
      (:ACTION ?varv)
      (sem (d::and ?semcnp ?semv)))
 -cnp->cnp-vpgerund- 0.9
 (head (cnp (var ?varcnp) (sem ?semcnp) (agr ?a) (gerund -) (nunit -)))
 (vp (vform prespart) (var ?varv) (sem ?semv) (aux -) (inv -) (modal -) (prep-lex (? !pl -))))

;; "Car that makes noise", "plant that employs 50 workers"
((cnp (var ?varn) (agr ?a) (modif-adj +)
      (:ACTION ?varvp)
      (:EVENT ?varvp)
      (:SUBJECT ?varn)
      (:NOUN ?varn)
      (sem (d::and ?semn ?semvp)))
 -cnp->cnp-that-clause-
 (head (cnp (var ?varn) (sem ?semn) (agr ?a) (modif-adj -)))
 (sconj (lex that))
 (vp (var ?varvp) (sem ?semvp) (vform ?vform) (agr ?a)))

;; "city where he is fighting"
((cnp (var ?varn) (agr ?a) (modif-adj +)
      (:ACTION ?varvp)
      (:EVENT ?varvp)
      (:SUBJECT ?varn)
      (:NOUN ?varn)
      (sem (d::and ?semn ?semvp (d::eventOccursAt ?varvp ?varn))))
 -cnp->n-where-clause-
 (head (noun (var ?varn) (sem ?semn) (agr ?a)))
 (sconj (lex where))
 (vp (inv +) (aux -) (link -) (agr ?a) (var ?varvp) (sem ?semvp)))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Noun Phrases - determination
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; "the cat", "a cat"
((np (var ?varcnp) (agr ?a) (det ?semdet) (globally-quantified ?gquant)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->det-cnp-
 (det (agr ?a) (var ?vardet) (sem ?semdet) (lex (? !l every some)) (wh -) (tposs -))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (gerund-dir-obj -) (globally-quantified
?gquant))))
```

```
;; conjunctive np
((np (var (d::ConjunctiveVar ?varnp1 ?varnp2))
     (agr (? a 2p 3p))
     (sem (d::and ?semnp1 ?semnp2)))
 -np->np-and-np-
 (head (np (var ?varnp1) (sem ?semnp1)))
 (cconj (lex and))
 (np (var ?varnp2) (sem ?semnp2)))

;; "what country"
((np (var ?varcnp) (agr ?a) (wh q) (stype d::WhatQuestion-IBT)
     (sem (d::thereExists ?varcnp (d::and ?semcnp (d::coreferent ?varcnp (d::WhQuestionFn
d::what))))))
 -np->what-cnp-
 (det (wh q) (lex what) (agr ?a))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

;; "which color"
((np (var ?varcnp) (agr ?a) (wh q) (stype d::WhichQuestion-IBT)
     (sem (d::thereExists ?varcnp (d::and ?semcnp (d::coreferent ?varcnp (d::WhQuestionFn
d::which))))))
 -np->which-cnp-
 (det (wh q) (lex which) (agr ?a))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

;;;;;;;;;;;;;;;;;;;;;;;;;
;;; undetermined nps

;; trick for setting the var field to a complex sem field
((np (var ?sem) (agr ?a))
 -np->np-set-var-sem-
 (head (np-svs (agr ?a) (sem ?sem))))

;; instantiated plural, as opposed to generic
;; plurals "cats"
;; either a group or the whole collection
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->cnp-plural-
 (head (cnp (agr (? a 1p 2p 3p)) (var ?varcnp) (sem ?semcnp) (gerund -))))


;; gerunds "running is good"
((np (var ?varcnp) (agr ?a) (gerund +)
     (sem ?semcnp))
 -np->cnp-gerund- 0.9
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (gerund +))))

;; collectives "there is blood", "buy bleach"
((np (var ?varcnp) (agr 3s)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->cnp-collective- 0.9
 (head (cnp (agr 3s) (var ?varcnp) (sem ?semcnp) (ncollective +))))

;; conceptual terms (catch all)
 ((np (var ?varcnp) (agr ?a)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->cnp-conceptual- 0.8
 (head (cnp (agr (? a 1s 2s 3s)) (var ?varcnp) (sem ?semcnp) (countable -) (gerund -) (ntime2 -)
(ntime1 -))))

;; proper names
```

```
((np (agr ?a) (var ?semn) (sem (d::properNameReference ?semn)))
 -np->pname- 0.9
 (head (pname (agr ?a) (var ?varn) (lex ?lexn) (orth ?orthn) (sem ?semn))))

((np (agr ?a) (var ?semn) (sem (d::properNameReference ?semn)))
 -np->the-pname- 0.9
 (det (lex the))
 (head (pname (agr ?a) (var ?varn) (lex ?lexn) (orth ?orthn) (sem ?semn))))

;;;;;;;;;;;;;;;;;
;; pronouns

;; "him", "her"
((np (var ?varpn) (agr ?a)
     (sem (d::thereExists ?varpn))
     (pro (d::PronounMappingFn ?varpn ?sempn)))
 -np->pronoun-
 (head (pronoun (agr ?a) (var ?varpn) (sem ?sempn) (wh -) (proposs -))))

;; who/what are pronouns
((np (var (d::WhQuestionFn d::who)) (agr ?a) (wh q) (stype d::WhoQuestion-IBT))
 -np->who-pronoun-
 (head (pronoun (agr ?a) (wh q) (lex who))))

((np (var (d::WhQuestionFn d::what)) (agr ?a) (wh q) (stype d::WhatQuestion-IBT))
 -np->what-pronoun-
 (head (pronoun (agr ?a) (wh q) (lex what))))

;;;;;;;;;;;;;;;;;
;; possessives

;; "his cat"
((np (var ?varcnp) (agr ?a)
     (pro (d::PronounMappingFn ?vardet ?semdet))
     (:POSSESSOR ?vardet)
     (sem (d::thereExists (d::TheList ?vardet ?varcnp)
                          (d::and ?semcnp
                                  (d::possessiveRelation ?vardet ?varcnp)))))
 -np->possessive-cnp-
 (det (agr ?a) (var ?vardet) (sem ?semdet) (tposs +))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

;; "the dove's life"
((np (var ?varcnp) (agr ?a)
     (:POSSESSOR ?varpos)
     (sem (d::thereExists ?varcnp
                          (d::and ?sempos
                                  ?semcnp
                                  (d::possessiveRelation ?varpos ?varcnp)))))
 -np->possessivenp-cnp-
 (np (agr ?a) (var ?varpos) (sem ?sempos))
 (misc (lex ^s))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

;;;;;;;;;;;;;;;;;
;; demonstrative references

((np (var (d::DemonstrativeFn ?lexp)) (agr (? a 1s 2s 3s 1p 2p 3p)))
 -np->dem-there-
 (head (adverb (lex (? lexp here there)) (loc&dir-adv +))))

;;;;;;;;;;;;;;;;;
;; wh-q pronoun as inverted np

;; "what I ate", "who he is"
((np (agr ?a) (var ?varwh) (inv-subject ?invsub)
     (:SUBJECT ?varwh)
     (:ACTION ?varvp)
```

```
     (sem (d::thereExists ?varwh ?semvp)))
 -np->whpn-vp-inv-
 (pronoun (agr ?a) (wh q) (sem ?semwh) (var ?varwh))
 (head (vp (inv +) (inv-subject ?invsub) (aux -) (link -) (agr ?a) (var ?varvp) (sem ?semvp))))

;; "whatever" isn't a (wh q)
((np (agr ?a) (var ?varwh) (inv-subject ?invsub)
     (:SUBJECT ?varwh)
     (:ACTION ?varvp)
     (sem (d::thereExists ?varwh ?semvp)))
 -np->whateverpn-vp-inv-
 (pronoun (agr ?a) (lex whatever) (sem ?semwh) (var ?varwh))
 (head (vp (inv +) (inv-subject ?invsub) (aux -) (link -) (agr ?a) (var ?varvp) (sem ?semvp))))

;; "whenever you ..."
((np (agr ?a) (var ?varwh) (inv-subject ?invsub)
     (:SUBJECT ?varwh)
     (:ACTION ?varvp)
     (sem (d::thereExists ?varwh (d::and (d::denotes ?varwh ?varvp) ?semvp))))
 -np->wheneverpn-vp-inv-
 (sconj (agr ?a) (lex whenever) (sem ?semwh) (var ?varwh))
 (head (vp (inv +) (inv-subject ?invsub) (aux -) (link -) (agr ?a) (var ?varvp) (sem ?semvp))))

;; "when it was", "where they went", "how he did it"
((np (agr ?a) (var ?varwh)
     (:SUBJECT ?varwh)
     (:ACTION ?varvp)
     (:LOCATION ?varwh)
     (:DATE ?varwh)
     (sem (d::thereExists ?varwh (d::and ?semvp ?semwh))))
 -np->whadv-vp-inv-
 (adverb (agr ?a) (wh q) (sem ?semwh) (var ?varwh))
 (head (vp (inv +) (aux -) (link -) (agr ?a) (var ?varvp) (sem ?semvp))))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Quantifying Groups
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; "one cat"
((np (var ?varcnp) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->one-cnp-
 (cardinal (var ?varc) (sem ?semc) (lex one))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr ?a))))

;; "1 cat"
((np (var ?varcnp) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->1-cnp-
 (number (var ?varnum) (sem ?semnum) (lex 1))
 (head (cnp (var ?varcnp) (sem ?semcnp) (nunit -) (agr ?a))))

;; "the one cat"
((np (var ?varcnp) (agr ?a) (det ?semdet)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->the-one-cnp-
 (det (agr ?a) (sem ?semdet) (lex the))
 (cardinal (var ?varc) (sem ?semc) (lex one))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr ?a))))

;; "the 1 cat"
((np (var ?varcnp) (agr ?a) (det ?semdet)
     (sem (d::thereExists ?varcnp ?semcnp)))
 -np->the-1-cnp-
 (det (agr ?a) (sem ?semdet) (lex the))
 (number (var ?varnum) (sem ?semnum) (lex 1))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr ?a))))
```

```
;; "one of the cats"
((np (var (:DEPSYM one-of- ?varpp)) (agr ?a) (det Definite-NLAttr)
     (sem (d::thereExists (d::TheList ?varpp (:DEPSYM one-of- ?varpp))
                          (d::and
                           (d::isa ?varpp d::Set-Mathematical)
                           (d::member (:DEPSYM one-of- ?varpp) ?varpp)
                           (d::forAll (:DEPSYM elt-of- ?varpp)
                                      (d::implies
                                       (d::member (:DEPSYM elt-of- ?varpp) ?varpp)
                                       (d::SublisFn ?varpp (:DEPSYM elt-of- ?varpp) ?sempp)))))))
 -np->one-of-np-
 (cardinal (var ?varc) (sem ?semc) (lex one))
 (head (pp (var ?varpp) (sem ?sempp) (gap -))))

;; "two cats"
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::cardinality (:DEPSYM group-of- ?varcnp) ?semc)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->card-cnp-
 (cardinal (var ?varc) (sem ?semc) (lex (? !lexc one)))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr (? a 1p 2p 3p)))))

;; "2 cats"
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::cardinality (:DEPSYM group-of- ?varcnp) ?semnum)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->num-cnp-
 (number (var ?varnum) (sem ?semnum) (lex (? !lexc 1)))
 (head (cnp (var ?varcnp) (sem ?semcnp) (nunit -) (agr (? a 1p 2p 3p)))))

;; "the two cats"
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det ?semdet)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::cardinality (:DEPSYM group-of- ?varcnp) ?semc)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->the-card-cnp-
 (det (agr ?a) (sem ?semdet) (lex the))
 (cardinal (var ?varc) (sem ?semc) (lex (? !lexc one)))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr (? a 1p 2p 3p)))))

;; "the 2 cats"
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det ?semdet)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::cardinality (:DEPSYM group-of- ?varcnp) ?semnum)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
```

```
 -np->the-num-cnp-
 (det (agr ?a) (sem ?semdet) (lex the))
 (number (var ?varnum) (sem ?semnum) (lex (? !lexc 1)))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp) (agr (? a 1p 2p 3p)))))

;; "many cats", "few times", "some days"
;; there are only 14 quantifiers in cyc, easier to handle here than build
;;  special handling in the connection to cyc
((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::qualitativeExtent (:DEPSYM group-of- ?varcnp) d::Many)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->many-cnp-
 (quant (lex many))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

((np (var (:DEPSYM group-of- ?varcnp)) (agr ?a) (det Indefinite-NLAttr)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::qualitativeExtent (:DEPSYM group-of- ?varcnp) d::Some)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->some-cnp-
 (quant (lex (? l some several)))
 (head (cnp (agr ?a) (var ?varcnp) (sem ?semcnp))))

;; likewise "some of the cats", "several of them"
((np (var (:DEPSYM subset-of- ?varnp)) (det Definite-NLAttr)
     (sem (d::thereExists (d::TheList ?varnp (:DEPSYM subset-of- ?varnp))
                          (d::and
                           (d::isa ?varnp d::Set-Mathematical)
                           (d::isa (:DEPSYM subset-of- ?varnp) d::Set-Mathematical)
                           (d::qualitativeExtent (:DEPSYM subset-of- ?varnp) d::Some)
                           (d::subsetOf (:DEPSYM subset-of- ?varnp) ?varnp)
                           (d::forAll (:DEPSYM elt-of- ?varnp)
                                      (d::implies
                                       (d::member (:DEPSYM elt-of- ?varnp) ?varnp)
                                       (d::SublisFn ?varnp (:DEPSYM elt-of- ?varnp) ?sempp)))))))
 -np->some-of-np-
 (quant (lex (? l some several)))
 (head (pp (var ?varnp) (sem ?sempp) (prep-lex of) (gap -))))

;;; "all dogs", "all his dogs", "all the dogs at the park"
((np (var ?varnp) (agr ?a)
     (sem (d::forAll ?varnp (d::implies ?semnp :SCOPED-CLAUSE))))
 -np->all-np-
 (quant (lex all))
 (head (np (agr ?a) (var ?varnp) (sem ?semnp))))

;;; JLT "a few" as in "a few months"
((np (var (:DEPSYM group-of- ?varcnp)) (arg ?a)
     (sem (d::thereExists (:DEPSYM group-of- ?varcnp)
                          (d::and
                           (d::isa (:DEPSYM group-of- ?varcnp) d::Set-Mathematical)
                           (d::qualitativeExtent (:DEPSYM group-of- ?varcnp) d::AFew-Quant)
                           (d::forAll ?varcnp
                                      (d::implies
                                       (d::member ?varcnp (:DEPSYM group-of- ?varcnp))
                                       ?semcnp))))))
 -np->a-few-cnp-
```

```
 (noun (lex a-few))
 (head (cnp (nunit +) (var ?varcnp) (agr ?a) (sem ?semcnp))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Unknown as label
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;;;; "cylinder C1"
((cnp (var ?semu) (agr ?a)
      (:NOUN ?semu)
      (sem ?semn))
 -cnp->noun-unknown- 0.8
 (head (noun (var ?varn) (sem ?semn) (agr ?a)))
 (unknown (var ?varu) (sem ?semu)))

;;;;; "cylinder C1" (undetermined)
((np (var ?semu) (agr ?a)
     (:NOUN ?semu)
     (sem (d::thereExists ?semu ?semn)))
 -np->noun-unknown- 0.8
 (head (noun (var ?varn) (sem ?semn) (agr ?a)))
 (unknown (var ?varu) (sem ?semu)))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Measure Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;;;; "one foot" "16 pounds" "23 inches per second"
;;;;; agreement is all jacked up: "a 2 pound fish" vs. "the fish is 2 pounds"
;;;;; could be split into two types of measure phrases
((mp (var ?varn) (agr ?a)
     (sem (?semn ?semnum)))
 -mp->number-unit-
 (number (sem ?semnum))
 (np-svs (nunit +) (var ?varn) (agr ?a) (sem ?semn)))

((np-svs (var ?varcnp) (agr ?a) (nunit +)
         (sem (d::ConceptForInstanceFn ?varcnp ?semcnp)))
 -np-svs->cnp-unit-
 (head (cnp (nunit +) (var ?varcnp) (agr ?a) (sem ?semcnp))))

;;; "13 years old", "2 inches long"
((adjp (var ?varmp)
       (:MEASURE ?semmp)
       (sem ?sema))
 -adjp->mp-adj-
 (mp (var ?varmp) (sem ?semmp))
 (head (adjective (var ?vara) (sem ?sema))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Verb Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; present, past, present participle, past participle
;; "eat", "ate", "eating", "eaten"
((vp (agr ?a) (vform ?vform) (var ?varv) (mwp ?mwp)
     (:ACTION ?varv)
     (sem (d::thereExists ?varv ?semv)))
 -vp->v-
 (head (verb (lex (? !l will)) (agr ?a) (vform (? vform pres past prespart pastpart))
             (mwp ?mwp) (var ?varv) (sem ?semv))))

;;;;;;;;;;;;;;;;;
;;; objects

;; vp + direct object "eat the fruit"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj +) (inv-subject ?invsub)
     (:ACTION ?varv)
```

```
     (:OBJECT ?varnp)
     (sem (d::and ?semnp ?semv)))
 -vp->vp-np-
 (head (vp (vform ?vform) (agr ?a) (var ?varv) (sem ?semv) (subcat (? s np np-pp part-np))
           (aux -) (modal -) (negated -) (inv -) (dir-obj -) (link -) (prep-lex -)))
 (np (var ?varnp) (sem ?semnp) (inf -) (ntime1 -) (inv-subject ?invsub)))

;; vp + direct object (mp) "move 1 inch"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj +) (inv-subject ?invsub)
     (:ACTION ?varv)
     (:MEASURE ?semmp)
     (sem ?semv))
 -vp->vp-mp-
 (head (vp (vform ?vform) (agr ?a) (var ?varv) (sem ?semv) (subcat (? s np np-pp part-np))
           (aux -) (modal -) (negated -) (inv -) (dir-obj -) (link -) (prep-lex -)))
 (mp (var ?varmp) (sem ?semmp)))

;; parenthetical vp + direct object "expect rain"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj +) (inv-subject ?invsub)
     (:ACTION ?varv)
     (:CLAUSE ?semnp)
     (sem ?semv))
 -vp->parentheticalvp-np-
 (head (vp (vform ?vform) (agr ?a) (var ?varv) (sem ?semv) (subcat (? s np np-pp part-np))
           (aux -) (modal -) (negated -) (inv -) (dir-obj -) (link -) (parenthetical +)))
 (np (var ?varnp) (sem ?semnp) (inf -) (ntime1 -) (inv-subject ?invsub)))

;; inverted ditransitive "to" frame: "gave him the book", "wrote her a letter"
((vp (var ?varv) (vform ?vform) (agr ?a)
     (:ACTION ?varv)
     (:OBJECT ?varnp2)
     (:OBLIQUE-OBJECT ?varnp1)
     (sem (d::and ?semv ?semnp)))
 -vp->vp-np-np-
 (head (vp (subcat (? s np-to-np)) (vform ?vform) (agr ?a) (var ?varv) (sem ?semv)
           (aux -) (modal -) (negated -) (inv -)))
 (np (var ?varnp1) (sem ?semnp1) (inf -) (ntime1 -))
 (np (var ?varnp2) (sem ?semnp2) (inf -) (ntime1 -)))


;;;;;;;;;;;;;;;;
;;; complements

;;; vp + prep "walking in an orchard"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj ?dir-obj)
    (:ACTION ?varv)
    (:OBLIQUE-OBJECT ?varpp)
    (:NOUN :SUBJECT)
    (prep-lex ?!lexp)
    (sem (d::and ?semv ?sempp)))
 -vp->vp-pp-
 (head (vp (subcat (? s pp np np-pp part-pp pp-pp)) (vform ?vform) (var ?varv) (agr ?a) (sem
?semv)
           (aux -) (modal -) (negated -) (inv -) (vpinf -) (dir-obj (? !dir-obj invalid))))
 (pp (var ?varpp) (sem ?sempp) (prep-lex (? !lexp of to from))))

;;; vmotion + to/from "move to the left"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid)
    (:ACTION ?varv)
    (:OBLIQUE-OBJECT ?varpp)
    (prep-lex ?plex)
    (sem (d::and ?semv ?sempp)))
 -vp->vp-vmotion-pp-
 (head (vp (vform ?vform) (var ?varv) (agr ?a) (sem ?semv) ;;(subcat (? s pp np-pp part-pp pp-
pp))
           (aux -) (modal -) (negated -) (inv -) (vpinf -) (vmotion +))) ;;(dir-obj -)
 (pp (var ?varpp) (sem ?sempp) (prep-lex (? plex to from))))
```

```
;;; unlikely non-vmotion verb + to "return to"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid)
    (:ACTION ?varv)
    (:OBLIQUE-OBJECT ?varpp)
    (prep-lex ?plex)
    (sem (d::and ?semv ?sempp)))
 -vp->vp-unlikely-pp- 0.8
 (head (vp (vform ?vform) (var ?varv) (agr ?a) (sem ?semv) ;;(subcat (? s pp np-pp part-pp pp-
pp))
          (aux -) (modal -) (negated -) (inv -) (vpinf -))) ;;(dir-obj -)
 (pp (var ?varpp) (sem ?sempp) (prep-lex (? plex to from))))

;;; vp + to (purpose) + vp (gerund) "ate to survive"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid)
     (:ACTION ?varv)
     (prep-lex to)
     (sem (d::and ?semv (d::purposeInEvent :SUBJECT ?varv ?semnp))))
 -vp->vp-to-npgerund- 0.9
 (head (vp (vform ?vform) (var ?varv) (agr ?a) (sem ?semv) ;;(subcat (? s pp np-pp part-pp pp-
pp))
          (aux -) (modal -) (negated -) (inv -) (dir-obj ?dir-obj) (vpinf -) (link -)))
 (np (var ?varnp) (sem ?semnp) (inf +) (np-inf -)))

;;; parenthetical vp + that + slp: "He believes that John ate the cake."
((vp (var ?varvp) (vform ?vform) (agr ?a) (dir-obj invalid)
    (:ACTION ?varvp)
    (:CLAUSE ?semslp)
    (sem ?semvp))
 -vp->vp-that-slp- 0.9
 (head (vp (vform ?vform) (var ?varvp) (agr ?a) (sem ?semvp) (parenthetical +)
          (aux -) (modal -) (negated -) (inv -)))
 (sconj (lex that))
 (slp (var ?varslp) (sem ?semslp)))

;;; parenthetical vp + inv "saw the dog running"
((vp (var ?varvp) (vform ?vform) (agr ?a) (dir-obj invalid)
    (:ACTION ?varvp)
    (:CLAUSE ?semvp2)
    (sem ?semvp))
 -vp->vp-vpinv-
 (head (vp (vform ?vform) (var ?varvp) (agr ?a) (sem ?semvp) (parenthetical +)
          (aux -) (modal -) (negated -) (inv -) (dir-obj -)))
 (vp (var ?varvp2) (sem ?semvp2) (inv +)))

;; subordinate if-clause "discover if this is true", "know if he came by"
((vp (var ?varvp) (vform ?vform) (agr ?a) (dir-obj invalid)
    (:ACTION ?varvp)
    (:CLAUSE ?semslp)
    (sem ?semvp))
 -vp->vp-if-slp- 0.9
 (head (vp (vform ?vform) (var ?varvp) (agr ?a) (sem ?semvp)
          (subcat pp-how-to-inf) (aux -) (modal -) (negated -) (inv -)))
 (sconj (lex if))
 (slp (var ?varslp) (sem ?semslp)))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Auxilliary Verbs
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;;;;;;;;;;;;;;;;
;;; aspects

;; progressive aspect "am/is/are eating", "was/were eating", "(will) be eating"
((vp (agr ?a) (vform ?auxform) (var ?varv) (aux +) (negated ?neg) (inv ?inv) (progressive +)
      (sem ?semv))
 -vp->vbe-vp-
 (aux (vbe +) (sem ?semaux) (var ?varaux) (agr ?a) (vform (? auxform past pres pastpart)))
 (head (vp (vform prespart) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))
```

```
;; perfect aspect "have eaten", "have been eating"
((vp (agr (? a 1p 2p 3p 1s 2s)) (vform pres) (var ?varv) (aux +) (negated ?neg) (inv ?inv)
(perfect +)
      (sem ?semv))
 -vp->have-vp-
 (aux (lex have) (sem ?semaux) (var ?varaux))
 (head (vp (vform pastpart) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;; perfect aspect "has eaten", "has been eating"
((vp (agr 3s) (vform pres) (var ?varv) (aux +) (negated ?neg) (inv ?inv) (perfect +)
      (sem ?semv))
 -vp->has-vp-
 (aux (lex has) (sem ?semaux) (var ?varaux))
 (head (vp (vform pastpart) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;; perfect aspect "had eaten", "had been eating"
((vp (agr ?a) (vform past) (var ?varv) (aux +) (negated ?neg) (inv ?inv) (perfect +)
      (sem ?semv))
 -vp->had-vp-
 (aux (lex had) (sem ?semaux) (var ?varaux) (agr ?a))
 (head (vp (vform pastpart) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;; perfect aspect "having eaten", "having been eating"
((vp (agr ?a) (vform past) (var ?varv) (aux +) (negated ?neg) (inv ?inv) (perfect +)
      (sem ?semv))
 -vp->having-vp-
 (aux (lex having) (sem ?semaux) (var ?varaux) (agr ?a))
 (head (vp (vform pastpart) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;;;;;;;;;;;;;;;;;
;;; future tense

;; future tense "will eat", "will be eating", "will have eaten", "will have been eating"
((vp (agr ?a) (vform future) (var ?varv) (modal +) (negated ?neg) (inv ?inv)
     (sem (d::willBe ?semv)))
 -vp->will-vp-
 (aux (modal +) (lex will)  (sem ?semaux) (var ?varaux) (agr ?a) (vform pres))
 (head (vp (vform pres) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;; future tense "going to eat"
((vp (agr ?a) (vform prespart) (var ?varv) (modal +) (negated ?neg) (inv ?inv)
     (sem (d::willBe ?semv)))
 -vp->going-to-vp-
 (verb (lex going))
 (prep (lex to))
 (head (vp (vform pres) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv))))

;;;;;;;;;;;;;;;;;
;;; modals

;; case for non-negated vp
((vp (agr ?a) (vform ?vform) (var ?varv) (modal +) (negated -) (inv ?inv) (link ?link)
     (:CLAUSE ?semv)
     (sem ?semaux))
 -vp->modal-vp-
 (aux (modal +) (lex (? !l did do does will)) (sem ?semaux) (var ?varaux) (agr ?a)
      (vform (? vform pres past)))
 (head (vp (vform pres) (var ?varv) (sem ?semv) (negated -) (inv ?inv) (link ?link))))

;; case for negated vp (need to invert not/modal in the semantics)
((vp (agr ?a) (vform ?vform) (var ?varv) (modal +) (negated ?neg) (inv ?inv) (link ?link)
     (:CLAUSE ?semv)
     (sem (d::not ?semaux)))
 -vp->modal-not-vp-
 (aux (modal +) (lex (? !l did do does will)) (sem ?semaux) (var ?varaux) (agr ?a) (vform
?vform))
 (adverb (lex not))
```

```
 (head (vp (vform pres) (var ?varv) (sem ?semv) (negated ?neg) (inv ?inv) (link ?link))))

;; did/do are empty modals
((vp (agr ?a) (vform ?vform) (var ?varv) (negated ?neg) (inv ?inv) (aux +)
     (sem ?semv))
 -vp->did-vp-
 (aux (modal +) (lex (? l did do does)) (vform (? vform past pres)) (agr ?a))
 (head (vp (vform pres) (var ?varv) (sem ?semv) (modal -) (negated ?neg) (inv ?inv))))

;; "did not", "does not" with a verb gap
((vp (agr ?a) (vform ?vform) (var ?varv) (negated +) (aux +)
     (sem (d::not (d::thereExists ?varv ?semv))))
 -vp->did-not-
 (head (verb (lex (? l did do does)) (var ?varv) (sem ?semv)))
 (adverb (lex not)))

;;;;;;;;;;;;;;;;;
;;; negation

;; negating a verb
((vp (agr ?a) (vform ?vform) (var ?varv) (negated +)
     (sem (d::not ?semv)))
 -vp->not-vp-
 (adverb (lex not))
 (head (vp (vform ?vform) (agr ?a) (var ?varv) (sem ?semv) (modal -))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Participles
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; participles act as adjectives, modifies the subject
;; "is capped", "is rotting", "damaged house"

 ((adjp (var ?varv) (participle +)
       (:OBJECT :NOUN)
       (:SUBJECT d::UnspecifiedPassive)
       (:ACTION ?varv)
       (sem ?semv))
 -adjp->vp- 0.9
 (head (vp (vform (? vform prespart pastpart)) (var ?varv) (sem ?semv)
           (:OBLIQUE-OBJECT ?oblique)
           (aux -) (modal -) (negated -) (inv -) (dir-obj -) (utterance -))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Infinitives
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; infinitive vp as noun phrase "to eat", "to run", "to die"
((np (var ?varv) (sem ?semv) (inf +)) ;; (inv-subject ?invsub))
 -np->to-vp-
 (prep (lex to))
 (head (vp (vform pres) (var ?varv) (sem ?semv) (agr ?a) (inv -)))) ;;(inv-subject ?invsub))))

;; vp + infinitive "wanted to run around the lake"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid) (vpinf +)
     (:ACTION ?varv)
     (:INF-COMP ?semnp)
     (sem ?semv))
 -vp->vp-npinf-
 (head (vp (vform ?vform) (subcat (? sc to-inf-sc to-inf-rs)) (agr ?a) (var ?varv) (sem ?semv)
           (dir-obj -) (modal -) (inv -) (vpinf -) (aux -) (vmotion -)))
 (np (var ?varnp) (sem ?semnp) (inf +) (np-inf -)))

;; vp + np + infinitive "wanted John to run around the lake"
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid) (vpinf +)
     (:ACTION ?varv)
     (:INF-COMP ?semnp)
     (:OBJECT ?varnp-obj)
```

```
     (sem ?semv))
 -vp->vp-np-npinf-
 (head (vp (vform ?vform) (subcat (? sc np-to-inf np-to-inf-vc np-to-inf-oc)) (agr ?a) (var
?varv) (sem ?semv)
           (modal -) (link -) (inv -) (aux -) (neg -) (mwp -) (vpinf -) (vmotion -)))
 (np (var ?varnp) (sem ?semnp) (inf +) (np-inf +) (np-obj ?varnp-obj)))

((np (var ?varinf) (agr ?a) (inf +) (np-inf +) (np-obj ?varnp)
      (:SUBJECT ?varnp)
      (:OBJECT :NO-OBJECT)
      (sem (d::and ?semnp ?seminf)))
 -np->np-npinf-
 (np (var ?varnp) (sem ?semnp) (agr ?a) (np-inf -))
 (head (np (var ?varinf) (sem ?seminf) (inf +) (np-inf -))))

;; way + infinitive "way to eat bananas" REM: other nouns besides "way"?
((cnp (var ?varn) (agr ?a)
      (:INF-COMP ?semnp)
      (sem ?semn))
 -cnp->way-npinf-
 (head (noun (agr ?a) (var ?varn) (sem ?semn) (lex way)))
 (np (var ?varnp) (sem ?semnp) (inf +)))

;; conditional infinitive
;; "promised to help him if he would swim"
;; (note that the promising is not conditional, which would be slp-sconj-slp)
((vp (var ?varv) (vform ?vform) (agr ?a) (dir-obj invalid) (vpinf +)
     (:ACTION ?varv)
     (:INF-COMP ?semnp)
     (:CLAUSE ?semslp)
     (sem ?semv))
 -vp->vp-npinf-conditional-
 (head (vp (vform ?vform) (subcat (? sc to-inf-sc to-inf-rs)) (agr ?a) (var ?varv) (sem ?semv)
           (dir-obj -) (modal -) (vpinf -) (vmotion -)))
 (np (var ?varnp) (sem ?semnp) (inf +) (np-inf -))
 (sconj (lex if))
 (slp (var ?varslp) (sem ?semslp)))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Gerunds
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; gerund as common noun phrase
((cnp (var ?varv) (agr (? a 1s 2s 3s 1p 2p 3p)) (gerund +) (gerund-dir-obj ?dir-obj)
      (:SUBJECT (:GAP :SUBJECT))
      (:OBJECT :OBLIQUE-OBJECT)
      (:ACTION ?varv)
      (sem ?semv))
 -cnp->vpgerund-
 (head (vp (vform prespart) (var ?varv) (sem ?semv) (aux -) (inv -) (modal -) (dir-obj ?dir-
obj))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Adverb Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; "newly" "fast"
((advp (var ?vara)
       (:GENERIC-VALUE-FN d::MediumToHighAmountFn)
       (sem ?sema))
 -advp->adverb-
 (head (adverb (var ?vara) (sem ?sema) (lex (? !l here there not)) (modif clausal-adv))))

;;; "very quickly"
((advp (var ?vara2)
       (:GENERIC-VALUE-FN (d::ConceptForInstanceFn ?sema1))
       (sem ?sema2))
```

```
 -advp->adverb-adverb-
 (adverb (var ?vara1) (sem ?sema1) (modif pre-adv))
 (head (adverb (var ?vara2) (sem ?sema2) (modif clausal-adv))))

;; coordinating conjunctions
;; REM: need to be much better constrained
((advp (var ?varc)
       (sem (d::and ?sema1 ?sema2)))
 -advp->advp-cconj-advp-
 (head (advp (var ?vara1) (sem ?sema1)))
 (cconj (var ?varc) (lex ?lexc))
 (advp (var ?vara2) (sem ?sema2)))

;;; advp + verb "quickly ran" "loudly said"
((vp (agr ?a) (vform ?vform) (var ?varv)
     (:ACTION ?varv)
     (:NOUN :SUBJECT)
     (sem (d::thereExists ?varv (d::and ?sema ?semv))))
 -vp->advp-v- 0.9
 (advp (sem ?sema) (modif clausal-adv))
 (head (verb (lex (? !l will)) (agr ?a) (vform (? vform base pres past prespart pastpart)) (var
?varv) (sem ?semv))))

;;; verb + advp "wrote very quickly"
((vp (agr ?a) (vform ?vform) (var ?varv)
     (:ACTION ?varv)
     (:NOUN :SUBJECT)
     (sem (d::thereExists ?varv (d::and ?sema ?semv))))
 -vp->v-advp- 0.9
 (head (verb (lex (? !l will)) (agr ?a) (vform (? vform base pres past prespart pastpart)) (var
?varv) (sem ?semv)))
 (advp (sem ?sema)))

;;; verb + advpart "went up", "looked down"
((vp (agr ?a) (vform ?vform) (var ?varv)
     (sem (d::thereExists ?varv (d::and ?sema ?semv))))
 -vp->v-advpart- 0.9
 (head (verb (lex (? !l will)) (agr ?a) (vform (? vform base pres past prespart pastpart))
             (subcat (? !subcat object-be)) (var ?varv) (sem ?semv)))
 (advpart (sem ?sema)))

((vp (agr ?a) (vform ?vform) (var ?varv) (link ?varv)
     (sem (d::thereExists ?varv ?sema)))
 -vp->v-advpart-object-be- 0.9
 (head (verb (agr ?a) (vform (? vform base pres past prespart pastpart))
             (subcat object-be) (var ?varv) (sem ?semv)))
 (advpart (sem ?sema)))

;;; verb + ntime "ate many times", "went yesterday"
((vp (agr ?a) (vform ?vform) (var ?varv)
     (:ACTION ?varv)
     (:NOUN :SUBJECT)
     (sem (d::and (d::thereExists ?varv ?semv)
                  ?semnp)))
 -vp->v-ntime-
 (head (verb (lex (? !l will)) (agr ?a) (subcat (? !sc object-be)) (vform (? vform base pres past
prespart pastpart)) (var ?varv) (sem ?semv)))
 (np (var ?varnp) (sem ?semnp) (ntime1 +) (gap -)))

;;; "northeast of"
((advp (var ?vara)
       (:OBLIQUE-OBJECT ?varpp)
       (prep-lex ?lexp)
       (sem (d::and ?sema ?sempp)))
 -advp->adv-pp- 0.9
 (head (adverb (var ?vara) (sem ?sema)))
 (pp (var ?varpp) (sem ?sempp) (prep-lex ?lexp)))
```

```
;; need constraint?  try (manner-adv +) or (modif pre-comparative)

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Implicit utterance
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; "said/yelled/whispered/etc _"
((vp (var ?varvp) (vform ?vform) (agr ?a) (dir-obj invalid) (comp +)
     (:ACTION ?varvp)
     (:CLAUSE ?semslp)
     (sem ?semvp))
 -vp->say-slp- 0.9
 (head (vp (agr ?a) (vsay +) (vform ?vform) (var ?varvp) (sem ?semvp)
           (aux -) (modal -) (negated -) (inv -) (comp -) (dir-obj -)))
 (slp (var ?varslp) (sem ?semslp)))

;; "said/yelled/whispered/etc that _"
((vp (var ?varvp) (vform ?vform) (agr ?a) (dir-obj invalid) (comp +)
     (:ACTION ?varvp)
     (:CLAUSE ?semslp)
     (sem ?semvp))
 -vp->say-that-slp-
 (head (vp (agr ?a) (vsay +) (vform ?vform) (var ?varvp) (sem ?semvp)
           (aux -) (modal -) (negated -) (inv -) (comp -)))
 (sconj (lex that))
 (slp (var ?varslp) (sem ?semslp)))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Explicit utterance subsentences
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; subsentence tokens
((s (var ?varpn) (sem ?sempn))
 -s->pname-
 (pname (sem ?sempn) (var ?varpn) (structural +)))

((s (var ?vars1) (sem (d::and ?sems1 ?sems2)))
 -s->s-s-
 (s (sem ?sems1) (var ?vars1))
 (s (sem ?sems2)))

;; "said/yelled/whispered/etc _"
((vp (var ?varvp) (vform ?vform) (agr ?a) (utterance +)
     (:CLAUSE ?sems)
     (:ACTION ?varvp)
     (sem ?semvp))
 -vp->said-comma-quote-s-quote-
 (head (vp (agr ?a) (vsay +) (vform (? vform pres past future prespart)) (var ?varvp) (sem
?semvp)
           (aux -) (modal -) (negated -) (inv -)))
 (punc (lex punc-comma))
 (punc (lex punc-double-quote))
 (s (sem ?sems))
 (punc (lex punc-double-quote)))

;; "said/yelled/whispered/etc <recipient> _"
((vp (var ?varvp) (vform ?vform) (agr ?a) (utterance +)
     (:CLAUSE ?sems)
     (:ACTION ?varvp)
     (:OBJECT ?varnp)
     (sem (d::and ?semnp ?semvp)))
 -vp->said-np-comma-quote-s-quote-
 (head (vp (agr ?a) (np-vsay +) (vform (? vform pres past future prespart)) (var ?varvp) (sem
?semvp)
           (aux -) (modal -) (negated -) (inv -)))
 (np (var ?varnp) (sem ?semnp))
 (punc (lex punc-comma))
 (punc (lex punc-double-quote))
```

```
 (s (sem ?sems))
 (punc (lex punc-double-quote)))

;; interjection utterance
((slp (var ?vari) (sem ?semi))
 -slp->interjection-
 (head (interjection (var ?vari) (sem ?semi))))

;; imperative utterance
((slp (var ?varvp) (vform ?vform) (stype d::ImperativeUtterance)
      (:SUBJECT (:GAP :SUBJECT))
      (sem ?semvp))
 -slp->vp-
 (head (vp (var ?varvp) (sem ?semvp) (vform (? vform pres)) (inv -))))

;; imperative utterance (polite)
((slp (var ?varslp) (stype d::ImperativeUtterance) (adjunct-adv +)
      (sem ?semslp))
 -slp->please-slp-
 (verb (lex please) (vform base))
 (head (slp (var ?varslp) (sem ?semslp) (stype d::ImperativeUtterance))))

;;; Inverted verb phrases

;;; "(did) the president order"
((vp (inv +) (agr ?a) (var ?varv) (vform ?vform) (utterance ?u)
     (:SUBJECT ?varnp)
     (:ACTION ?varv)
     (:OBJECT :SUBJECT)
     (:INF-COMP :SUBJECT)
     (inv-subject ?varnp)
     (sem (data::and ?semnp ?semv)))
 -vp->np-vp- 0.9
 (np (var ?varnp) (sem ?semnp) (inf -))
 (head (vp (vform ?vform) (var ?varv) (sem ?semv) (inv -) (utterance ?u))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Linking Verb "to be"
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; the link feature is used to constrain recursive complementing and for some
;;  higher level substitution patterns

;; adjective complement modifies the subject
;; "is green"
((vp (agr ?a) (vform ?vform) (var ?varvp) (link ?varvp)
     (:NOUN :SUBJECT)
     (:ACTION ?varvp)
     (sem (d::thereExists ?varvp ?semadjp)))
 -vp->be-adjp-
 (head (vp (agr ?a) (vform ?vform) (var ?varvp) (subcat object-be)
           (inv -) (link -) (modal -) (aux -) (negated -)))
 (adjp (sem ?semadjp) (var ?varadjp) (participle -)))

;; participle becomes head var
((vp (agr ?a) (vform ?vform) (var ?varadjp) (link ?varvp)
     (:NOUN :SUBJECT)
     (:ACTION ?varadjp)
     (sem (d::thereExists ?varvp ?semadjp)))
 -vp->be-participle-
 (head (vp (agr ?a) (vform ?vform) (var ?varvp) (subcat object-be) (inv -) (link -) (modal -)
(aux -)))
 (adjp (sem ?semadjp) (var ?varadjp) (participle +)))

;; subject complement reidentifies the subject
;; "is a fish"
((vp (agr ?a) (vform ?vform) (var ?varvp) (link ?varvp)
     (sem (d::and ?semnp
```

```
                 (d::denotes :SUBJECT ?varnp))))
 -vp->be-np- 0.9
 (head (vp (agr ?a) (vform ?vform) (var ?varvp) (subcat object-be)
          (link -) (aux -) (modal -) (negated -) (inv -)))
 (np (agr ?a) (sem ?semnp) (var ?varnp) (gerund -) (inf -)))

;;; be with 3s noun, requires modal (thus invalid agreement)
((vp (agr invalid) (vform ?vform) (var ?varvp) (link ?varvp)
     (sem (d::and ?semnp
                 (d::denotes :SUBJECT ?varnp))))
 -vp->be3s-np- 0.9
 (head (verb (vform ?vform) (var ?varvp) (subcat object-be) (lex be)))
 (np (agr 3s) (sem ?semnp) (var ?varnp) (gerund -) (inf -)))

;; np be infinitive complement
;; "(my dream) is to join the army"
;; SEF - Changed from coreferent-Underspecified
((vp (agr ?a) (vform ?vform) (var ?varvp) (link ?varvp)
     (:SUBJECT (:GAP :SUBJECT))
     (sem (d::denotes :SUBJECT ?seminf)))
 -vp->be-inf-
 (head (verb (agr ?a) (var ?varvp) (subcat object-be) (inv -) (link -)))
 (np (inf +) (np-inf -) (sem ?seminf) (gap -)))

;; prepositional complement
;; "is in the room"
;; is-pp has a different substitution pattern than vp-pp
((vp (var ?varv) (vform ?vform) (agr ?a) (link ?varv)
     (:NOUN :SUBJECT)
     (sem ?sempp))
 -vp->be-pp-
 (head (vp (agr ?a) (vform ?vform) (var ?varv) (subcat object-be) (inv -) (link -) (dir-obj -)))
 (pp (var ?varpp) (sem ?sempp)))

((vp (var ?varv) (vform ?vform) (agr ?a) (link ?varv)
     (:NOUN :SUBJECT)
     (sem (not ?sempp)))
 -vp->be-not-pp-
 (head (vp (agr ?a) (vform ?vform) (var ?varv) (subcat object-be) (inv -) (dir-obj -)))
 (adverb (lex not))
 (pp (var ?varpp) (sem ?sempp)))

;;; subject complement with measure phrase
;;; "is 3 kilograms"
((vp (agr ?agr) (vform ?vform) (var ?varv) (link ?varv)
     (sem (d::measure :SUBJECT ?semmp)))
 -vp->v-mp-object-be-
 (head (verb (agr ?agr) (vform ?vform) (var ?varv) (sem ?semv) (subcat object-be)))
 (mp (var ?varmp) (sem ?semmp)))

;;; "is 9 feet long"
((vp (agr ?agr) (vform ?vform) (var ?varv) (link ?varv)
     (:MEASURE ?semmp)
     (:NOUN :SUBJECT)
     (sem ?sema))
 -vp->v-mp-adj-object-be-
 (head (verb (agr ?agr) (vform ?vform) (var ?varv) (sem ?semv) (subcat object-be)))
 (mp (var ?varmp) (sem ?semmp))
 (adjective (var ?vara) (sem ?sema)))

;;; "is 9 feet from the water"
((vp (agr ?agr) (vform ?vform) (var ?varv) (link ?varv)
     (:NOUN :SUBJECT)
     (sem (d::and ?sempp (d::measure ?prep-var ?semmp))))
 -vp->v-mp-pp-object-be-
 (head (verb (agr ?agr) (vform ?vform) (var ?varv) (sem ?semv) (subcat object-be)))
 (mp (var ?varmp) (sem ?semmp))
 (pp (var ?varpp) (sem ?sempp) (prep-var ?prep-var)))
```

```
;;; "make the cat sick"
((vp (vform pres) (var ?varvp) (agr ?a)
     (:NOUN ?varnp)
     (:ACTION ?varvp)
     (:CLAUSE ?semadjp)
     (sem (d::thereExists ?varvp (d::and ?semnp ?semvp))))
 -vp->make-np-adjp-
 (head (verb (root make2) (var ?varvp) (sem ?semvp) (vform pres)))
 (np (var ?varnp) (sem ?semnp) (agr ?a))
 (adjp (sem ?semadjp) (var ?varadjp)))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Prepositional Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; "in an orchard"
((pp (var ?varnp)
     (prep-lex ?lexp)
     (prep-var ?varp)
     (:OBJECT ?varnp)
     (:POSSESSOR :NOUN)
     (sem (d::and ?semp ?semnp)))
 -pp->prep-np-
 (head (prep (var ?varp) (sem ?semp) (lex ?lexp)))
 (np (var ?varnp) (sem ?semnp) (np-inf -) (generic -)))

;; special case "out of"
((pp (var ?varnp)
     (prep-lex out-of)
     (prep-var ?varp)
     (:OBJECT ?varnp)
     (:POSSESSOR :NOUN)
     (sem (d::and ?semp ?semnp)))
 -pp->out-of-np-
 (head (prep (var ?varp) (sem ?semp) (lex out)))
 (prep (lex of))
 (np (var ?varnp) (sem ?semnp) (np-inf -) (generic -)))

;; instrumental
((pp (var ?varvp) (:OBJECT ?varvp)
     (prep-lex by)
     (sem (d::and ?semp ?semvp)))
 -pp->by-vp-
 (head (prep (var ?varp) (sem ?semp) (lex by)))
 (vp (var ?varvp) (sem ?semvp) (inv -) (subcat (? !subcat object-be))))

;; "either way" REM: adverbial, very similar to the ntime stuff
((pp (var ?varn)
     (:NOUN ?varn)
     (prep-lex either)
     (sem (d::and ?semp ?semn)))
 -pp->either-way-
 (head (det (var ?varp) (sem ?semp) (lex either)))
 (noun (var ?varn) (sem ?semn) (lex way)))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Adjective Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; "green", "big"
((adjp (var ?vara) (sem ?sema))
 -adjp->adj-
 (head (adjective (var ?vara) (sem ?sema))))

;;; "very green" "slightly dry"
((adjp (var ?varadj)
```

```
       (:GENERIC-VALUE-FN (d::ConceptForInstanceFn ?semadvp))
       (sem ?semadj))
 -adjp->adverb-adj- 0.9
 (adverb (modif pre-adj) (sem ?semadvp))
 (head (adjective (var ?varadj) (sem ?semadj))))

;;; "most ..." (comlex doesn't have this as pre-adj for some reason)
((adjp (var ?varadj) (global-quantifier +)
       (sem (d::not (d::thereExists ?varadv
                             (d::greaterThanOrEqualTo ((d::ConceptForInstanceFn ?semadj)
?varadv)
                                                      ((d::ConceptForInstanceFn ?semadj)
:NOUN))))))
 -adjp->most-adj-
 (adverb (lex most) (var ?varadv))
 (head (adjective (var ?varadj) (sem ?semadj))))

;;; "green with envy"
((adjp (var ?vara)
       (:OBLIQUE-OBJECT ?varpp)
       (prep-lex ?lexp)
       (sem (d::and ?sema ?sempp)))
 -adjp->adj-pp- 0.9
 (head (adjective (var ?vara) (sem ?sema)))
 (pp (var ?varpp) (sem ?sempp) (prep-lex ?lexp)))

;; adj + that phrase "sure that he is stupid"
((adjp (var ?varvp) (vform ?vform) (agr ?a)
    (:CLAUSE ?semslp)
    (sem ?sema))
 -vp->adj-that-slp- 0.9
 (head (adjective (var ?vara) (sem ?sema) (subcat (? sc that-s-adj extrap-adj-that-s))))
 (sconj (lex that))
 (slp (var ?varslp) (sem ?semslp)))

;; coordinating conjunctions
;; REM: but needs some kind of clausal handling
((adjp (var ?varc)
       (sem (d::and ?sema1 ?sema2)))
 -adjp->adjp-cconj-adjp-
 (head (adjp (var ?vara1) (sem ?sema1)))
 (cconj (var ?varc) (lex ?lexc))
 (adjp (var ?vara2) (sem ?sema2)))

;; adj + infinitive (w/o inverted subject)
((adjp (var ?varinf) (inf +)
       (:NOUN ?varinf)
       (:SUBJECT (:GAP :SUBJECT))
       (:OBJECT :SUBJECT)
       (sem (d::and ?sema ?seminf)))
 -adjp->adjp-npinf- 0.9
 (adjp (var ?vara) (sem ?sema))
 (head (np (var ?varinf) (sem ?seminf) (agr ?a) (inf +) (inv-subject -))))

;; adj + infinitive (w/ inverted subject)
((adjp (var ?varinf) (inf +)
       (:NOUN ?varinf)
       (:SUBJECT ?!invsub)
       (:OBJECT :SUBJECT)
       (sem (d::and ?sema ?seminf)))
 -adjp->adjp-npinf-invsub- 0.9
 (adjp (var ?vara) (sem ?sema))
 (head (np (var ?varinf) (sem ?seminf) (agr ?a) (inf +) (inv-subject (? !invsub -)))))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Sentence Level Phrases
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
;; normal, positive case
((slp (agr ?a) (var ?varvp) (stype ?st) (utterance ?u)
      (:ACTION ?varvp) (:EVENT ?varvp) (:SUBJECT ?varnp)
      (sem (d::and ?semnp ?semvp)))
 -slp->np-vp-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (head (vp (inv -) (link -) (subcat (? !subcat object-be))
           (negated -) (agr ?a) (var ?varvp) (sem ?semvp) (utterance ?u))))

;; negated case requires an aux or a modal
((slp (agr ?a) (var ?varvp) (stype ?st) (utterance ?u)
      (:ACTION ?varvp) (:EVENT ?varvp) (:SUBJECT ?varnp)
      (sem (d::and ?semvp ?semnp)))
 -slp->np-vp-negaux-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (head (vp (inv -) (link -) (subcat (? !subcat object-be))
           (negated +) (aux +) (agr ?a) (var ?varvp) (sem ?semvp) (utterance ?u))))

((slp (agr ?a) (var ?varvp) (stype ?st) (utterance ?u)
      (:ACTION ?varvp) (:EVENT ?varvp) (:SUBJECT ?varnp)
      (sem (d::and ?semvp ?semnp)))
 -slp->np-vp-negmodal-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (head (vp (inv -) (link -) (subcat (? !subcat object-be))
           (negated +) (aux -) (modal +) (agr ?a) (var ?varvp) (sem ?semvp) (utterance ?u))))

;; an inverted vp requires a wh-q np
((slp (agr ?a) (var ?varvp) (stype ?st)
      (:ACTION ?varvp) (:EVENT ?varvp) (:SUBJECT ?varnp)
      (sem (d::and ?semvp ?semnp)))
 -slp->np-vp-inv-
 (np (agr ?a) (wh q) (sem ?semnp) (var ?varnp) (stype ?st))
 (head (vp (inv +) (aux +) (link -) (subcat (? !subcat object-be))
           (agr ?a) (var ?varvp) (sem ?semvp) (inv-subject ?subvp))))

;;; special case for linking verbs
;; REM: why didn't I want action in there?  What does it break?
;; how else can I get the scope?
((slp (agr ?a) (var ?varvp) (stype ?st)
      (:SUBJECT ?varnp)
;;      (:ACTION ?varvp)
      (sem (d::and ?semnp ?semvp)))
 -slp->np-vp-link-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (head (vp (inv -) (link (? !vlink -)) (agr ?a) (var ?varvp) (sem ?semvp))))

;;; slp, vp-progressive
;;; CONTROL: subject could be ?subslp as well
((slp (var ?varslp) (stype ?st)
      (:SUBJECT ?varslp)
      (sem (d::and ?semslp ?semvp)))
 -slp->slp-comma-vp-
 (head (slp (var ?varslp) (sem ?semslp) (agr ?a) (stype ?st) (:SUBJECT ?subslp)))
 (punc (lex punc-comma))
 (vp (var ?varvp) (sem ?semvp) (vform prespart)))

;;; progressive verb aside
((slp (agr ?a) (var ?varvp2) (stype ?st) (utterance ?u) (aside ?varvp1)
      (:SUBJECT ?varnp)
      (sem (d::and ?semnp ?semvp1 ?semvp2
                   (d::temporallyIntersects ?varvp1 ?varvp2))))
 -slp->np-comma-vpprogressive-comma-vp-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (punc (lex punc-comma))
 (vp (inv -) (link -) (vform prespart) (var ?varvp1) (sem ?semvp1))
 (punc (lex punc-comma))
 (head (vp (inv -) (link -) (negated -) (agr ?a) (var ?varvp2) (sem ?semvp2) (utterance ?u))))
```

```
;;; progressive perfect verb aside
((slp (agr ?a) (var ?varvp2) (stype ?st) (utterance ?u) (aside ?varvp1)
      (:SUBJECT ?varnp)
      (sem (d::and ?semnp ?semvp1 ?semvp2
                   (d::after ?varvp2 ?varvp1))))
 -slp->np-comma-vpperfect-comma-vp-
 (np (agr ?a) (sem ?semnp) (var ?varnp) (stype ?st))
 (punc (lex punc-comma))
 (vp (inv -) (link -) (vform past) (perfect +) (var ?varvp1) (sem ?semvp1))
 (punc (lex punc-comma))
 (head (vp (inv -) (link -) (negated -) (agr ?a) (var ?varvp2) (sem ?semvp2) (utterance ?u))))

;;; preceeding pp
((slp (var ?varslp) (stype ?!stype)
      (:ACTION ?varslp)
      (sem (d::and ?semslp ?sempp)))
 -slp->pp-comma-slp-
 (pp (var ?varpp) (sem ?sempp) (prep-lex ?lexp))
 (punc (lex punc-comma))
 (head (slp (var ?varslp) (sem ?semslp) (inv -) (stype (? !stype d::ImperativeUtterance)))))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Sentence Adjunct Adverbials

;; leading clausal adverbs
;; meta-adv and temporal-adv seem connected here
((slp (var ?varslp)
      (adjunct-adv ?lexa)
      (sem ?semslp))
 -slp->adjunct-adverbial-comma-slp-
 (adverb (var ?vara) (sem ?sema) (lex ?lexa) (modif clausal-adv))
 (punc (lex punc-comma))
 (head (slp (var ?varslp) (sem ?semslp) (inv -) (stype (? !stype d::ImperativeUtterance)))))

((slp (var ?varslp)
      (adjunct-adv ?lexa)
      (sem ?semslp))
 -slp->adjunct-adverbial-slp- 0.8
 (adverb (var ?vara) (sem ?sema) (lex ?lexa) (modif clausal-adv))
 (head (slp (var ?varslp) (sem ?semslp) (inv -) (stype (? !stype d::ImperativeUtterance)))))

;; leading cconj
((slp (var ?varslp)
      (adjunct-adv ?lexcconj)
      (sem ?semslp))
 -slp->leading-cconj-comma-slp-
 (cconj (var ?varcconj) (sem ?semcconj) (lex ?lexcconj))
 (punc (lex punc-comma))
 (head (slp (var ?varslp) (sem ?semslp) (inv -) (stype (? !stype d::ImperativeUtterance)))))

;; leading coreference
;; SEF - changed from coreference-Underspecified
((slp (var ?varnp)
      (adjunct-adv +)
      (sem (d::and ?semnp ?semslp
                   (d::denotes ?varnp ?varslp))))
 -slp->np-be-comma-slp-
 (head (np (var ?varnp) (sem ?semnp) (agr ?a)))
 (verb (agr ?a) (vform ?vform) (var ?varv) (sem ?semv) (subcat object-be))
 (punc (lex punc-comma))
 (slp (var ?varslp) (sem ?semslp) (inv -) (stype (? !stype d::ImperativeUtterance))))

;; temporal noun phrase as sentence adjunct adverbial
;; ntime2 nouns may be used unadorned, "yesterday", "Friday", "midday" etc...
((slp (agr ?a) (var ?varslp)
      (sem (d::thereExists ?varn
```

```
            (d::and
             (d::temporallySubsumes ?varn ?varslp)
             ?semn ?semslp))))
 -slp->ntime2-slp-
 (noun (var ?varn) (sem ?semn) (ntime2 +))
 (head (slp (var ?varslp) (sem ?semslp))))

((slp (agr ?a) (var ?varslp)
      (sem (d::thereExists ?varn
            (d::and
             (d::temporallySubsumes ?varn ?varslp)
             ?semn ?semslp))))
 -slp->slp-ntime2-
 (head (slp (var ?varslp) (sem ?semslp)))
 (noun (var ?varn) (sem ?semn) (ntime2 +)))

;; ntime1 nouns require a timetag modifier
;; REM: determiners, timetag/ntime2, timetag alone?
((slp (agr ?a) (var ?varslp)
      (sem (d::thereExists ?varnp
            (d::and
             (d::temporallySubsumes ?varnp ?varslp)
             ?semnp ?semslp))))
 -slp->ntime1-phrase-slp-
 (np (var ?varnp) (sem ?semnp) (ntime1 +))
 (head (slp (var ?varslp) (sem ?semslp) (stype -))))

((slp (agr ?a) (var ?varslp)
      (sem (d::thereExists ?varnp
            (d::and
             (d::temporallySubsumes ?varnp ?varslp)
             ?semnp ?semslp))))
 -slp->slp-ntime1-phrase-
 (head (slp (var ?varslp) (sem ?semslp)))
 (np (var ?varnp) (sem ?semnp) (ntime1 +)))

;; timetag adjectives go before ntime1 nouns, "last June", "next hour"
((advp (agr ?a) (var ?varn) (ntime1 +)
       (:NOUN ?varn)
       (:REPLACE ?semn)
       (sem (d::and ?semn ?sema)))
 -advp->atimetag-ntime1-
 (adjective (var ?vara) (sem ?sema) (atimetag +))
 (head (noun (agr ?a) (var ?varn) (sem ?semn) (ntime1 +))))

;; "In return, ..."
((slp (var (sconj ?varslp in ?varav)) (sem ?semslp)
      (:SUBJECT ?subslp)
      (:ACTION ?varav)
      (:CLAUSE ?semslp)
      (sem (d::thereExists (d::TheList ?subslp ?varav) ?semav)))
 -vp->in-return-slp-
 (sconj (lex in))
 (verb (lex return) (vform base) (var ?varav) (sem ?semav))
 (punc (lex punc-comma))
 (head (slp (var ?varslp) (sem ?semslp) (:SUBJECT ?subslp))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Coordinating Conjunctions
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; coordinating conjunctions (with and without comma)
((slp (var (d::cconj ?varslp1 ?varslp2))
      (:ACTION ?varslp1)
      (sem (d::and ?semslp1 ?semslp2)))
 -slp->slp-cconj-slp-
 (head (slp (var ?varslp1) (sem ?semslp1) (adjunct-adv -)))
 (cconj (var ?varc) (sem ?semconj) (lex ?lexc))
```

```
 (slp (var ?varslp2) (sem ?semslp2) (stype (? !stype d::ImperativeUtterance))))

((slp (var (d::cconj ?varslp1 ?varslp2))
      (:ACTION ?varslp1)
      (sem (d::and ?semslp1 ?semslp2)))
 -slp->slp-comma-cconj-slp-
 (head (slp (var ?varslp1) (sem ?semslp1) (adjunct-adv -)))
 (punc (lex punc-comma))
 (cconj (var ?varc) (sem ?semconj) (lex ?lexc))
 (slp (var ?varslp2) (sem ?semslp2) (stype (? !stype d::ImperativeUtterance))))

;; coordinating conjunctions, without second subject (with and without comma)
((vp (var (d::cconj ?varvp1 ?varvp2))
     (sem (d::and ?semvp1 ?semvp2))
     (agr ?a)
     (vform ?vform)
     (aux +) (dir-obj invalid) (link ?link))
 -vp->vp-cconj-vp-
 (head (vp (var ?varvp1) (sem ?semvp1) (agr ?a) (vform (? vform pres past future prespart))
           (modal -) (link ?link)))
 (cconj (var ?varc) (sem ?semconj) (lex ?lexc))
 (vp (var ?varvp2) (sem ?semvp2) (agr ?a) (vform (? vform pres past future prespart))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Subordinating Conjunctions
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; They had proceeded a short distance when they met a lion.
((slp (var (d::sconj ?varslp1 d::when ?varslp2))
      (sem (d::and ?semslp1 ?semslp2
                   (d::after ?varslp2 ?varslp1))))
 -slp->slpperfect-sconj-slp-
 (slp (var ?varslp1) (sem ?semslp1) (perfect +))
 (sconj (lex when))
 (head (slp (var ?varslp2) (sem ?semslp2))))

;;; slp + so that (purpose) + slp "he got money so that he can buy a house"
((slp (var (d::sconj ?varslp d::so-that ?varslp2)) (stype ?stype)
      (sem (d::and ?semslp
                   (d::enables-Generic ?varslp ?semslp2))))
 -slp->slp-so-that-slp-
 (head (slp (var ?varslp) (sem ?semslp) (stype ?stype) (adjunct-adv -)))
 (sconj (lex so))
 (sconj (lex that))
 (slp (var ?varslp2) (sem ?semslp2) (stype (? !stype2 d::ImperativeUtterance))))

;;; While you were sleeping, they went home.
((slp (var ?varslp2)
      (sem (d::and (d::startsDuring ?varslp2 ?varslp1) ?semslp1 ?semslp2)))
 -slp->while-slp-comma-slp-
 (head (sconj (sem ?semconj) (var ?varconj) (lex while)))
 (slp (var ?varslp1) (sem ?semslp1))
 (punc (lex punc-comma))
 (slp (var ?varslp2) (sem ?semslp2)))

;;; the sconj "because" can also take an "of-NP" PP
;;; not sure how that generalizes, so treated specially for now

;;; Because of the rain, we stayed inside.
((slp (var ?varconj)
      (sem (d::and
            (d::causes-ThingSit ?varnp ?varslp2)
            ?semnp ?semslp2)))
 -slp->because-of-np-comma-slp-
 (head (sconj (sem ?semconj) (var ?varconj) (lex because)))
 (prep (lex of))
 (np (var ?varnp) (sem ?semnp))
 (punc (lex punc-comma))
```

```
 (slp (var ?varslp2) (sem ?semslp2) (vform (? !vform future))))

;;; future consequence
((slp (var ?varconj)
      (sem (d::and ?semnp (d::causes-ThingProp ?varnp ?semslp2))))
 -slp->because-of-np-comma-slp-future-
 (head (sconj (sem ?semconj) (var ?varconj) (lex because)))
 (prep (lex of))
 (np (var ?varnp) (sem ?semnp))
 (punc (lex punc-comma))
 (slp (var ?varslp2) (sem ?semslp2) (vform future)))

;;; We cancelled the game because of the heat.
((slp (var ?varslp1)
      (sem (d::and ?semslp ?semnp
                   (d::causes-ThingSit ?varnp ?varslp))))
 -slp->slp-because-of-np-
 (slp (var ?varslp) (sem ?semslp))
 (head (sconj (sem ?semconj) (var ?varconj) (lex because)))
 (prep (lex of))
 (np (var ?varnp) (sem ?semnp)))

;;; "If he ate poison, he will die."
((slp (var ?varslp)
      (sem (d::implies-DrsDrs ?semsubord ?semslp)))
 -slp->if-subordslp-comma-slp-
 (head (sconj (lex if) (sem ?semconj) (var ?varconj)))
 (slp (var ?varsubord) (sem ?semsubord))
 (punc (lex punc-comma))
 (slp (var ?varslp) (sem ?semslp)))

;;; "The opening would cause a fish to die."
;; REM: move to MWP if this works out
((slp (var ?varv)
      (sem (d::and ?semnp
                   (d::causes-Hypothetical
                    ?varnp (d::InterpretationOfClauseFn (d::ConstitNameByPosFn 4)))
                   (d::constitSubClause (d::ConstitNameByPosFn 4) ?semslp)
                   (d::discourseVarNLAttr ?varaux d::Structural)
                   (d::discourseVarNLAttr ?varv d::Structural))))
 -slp->np-would-cause-np-npinf-
 (head (np (var ?varnp) (sem ?semnp)))
 (aux (lex would) (var ?varaux) (sem ?semaux))
 (verb (lex cause) (var ?varv) (vform base) (sem ?semv))
 (slp (var ?varslp) (gap ?g) (sem ?semslp) (inf +)))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Sentences
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; stypes are kind of lame, but useful all the same

;;; pass up stype from the slp
((s (var ?vars)
    (stype ?!stype)
    (sem ?sems))
 -s->slp-stype-
 (head (slp (var ?vars) (sem ?sems) (stype (? !stype -))))
 (punc (lex (? l punc-period punc-exclamation-mark punc-question-mark))))

;; if the slp doesn't have one, use defaults
((s (var ?vars)
    (stype d::DeclarativeUtterance)
    (sem ?sems))
 -s->slp-declarative-
 (head (slp (var ?vars) (sem ?sems) (stype -)))
 (punc (lex (? l punc-period punc-exclamation-mark))))
```

```
((s (var ?vars)
    (stype d::InterrogativeUtterance)
    (sem ?sems))
 -s->slp-question-
 (head (slp (var ?vars) (sem ?sems) (stype -)))
 (punc (lex punc-question-mark)))

;; no handling for colon/semicolon
((s (var ?vars) (sem ?sems))
 -s->slp-colon-
 (head (slp (var ?vars) (sem ?sems)))
 (punc (lex (? l punc-colon punc-semicolon))))

;; quoted text can have the punctuation inside
((s (var ?vars)
    (stype d::DeclarativeUtterance)
    (sem ?sems))
 -s->slp-quoted-
 (head (slp (var ?vars) (sem ?sems) (utterance +))))

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; Generics & Defaults - SEF
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;;; "every dog"
((np (var ?varcnp) (agr ?a)
     (sem (d::forAll ?varcnp (d::implies ?semcnp :SCOPED-CLAUSE))))
 -np->every-cnp-
 (det (lex every))
 (head (cnp (agr (? a 1s 2s 3s)) (var ?varcnp) (sem ?semcnp))))

;;; "type of mammal"
((cnp (var ?varcnp) (agr ?a)
      (sem (d::refersToTypeOf ?varcnp (d::ConceptForInstanceFn ?semcnp))))
 -cnp->type-of-cnp-
 (noun (lex type-of))
 (head (cnp (agr (? a 1s 2s 3s)) (var ?varcnp) (sem ?semcnp))))

;; SEF - put here to do generics...
;; plurals "cats"
;; With this rule, plural denotes the type; not a group.
;; However, this doesn't work, because coreferent-Underspecified is symmetric,
;; so we can't really figure out which genls which...
((np (var ?varcnp) (agr ?a) (generic +)
     (sem (d::forAll ?varcnp (d::implies (d::and (d::variableInstance ?varcnp) ?semcnp) :SCOPED-
CLAUSE))))
;;(d::and (d::generalizeToTypeOf ?varcnp) ?semcnp)))
 -np->cnp-plural-generic-
 (head (cnp (agr (? a 1p 2p 3p)) (var ?varcnp) (sem ?semcnp) (gerund -))))

;;;;;;;;;;
;; Qualifiers
;;;;;;;;;;

;; 75 percent of the time, elephants are gray.
((slp (var ?varnp) (agr ?a) (adjunct-adv +) (stype ?stype)
      (sem (d::and ?semnp (d::ruleQualifier ?varnp ?semslp))))
 -slp->np-slp- 0.9
 (np (var ?varnp) (sem ?semnp) (ntime1 -))
 (punc (lex punc-comma))
 (head (slp (var ?varslp) (sem ?semslp) (stype ?stype))))


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; End of Code
```

8.2 Appendix B: Iranian folktales

Additional Iranian folktale variants from Dehghani et al. 2009.

Hossein Sacrifice		
Base	During the Iran and Iraq war, Hossein, a 13 year old boy who has sneaked into the army, is confronted with a convoy of tanks that if not stopped will destroy a part of the city that the boy is fighting at. Hossein can either try to run to his commander on time, inform him about the situation and save his own life or he can stop a tank by sacrificing his own life. Hossein, therefore, took a grenade from a nearby body, pulled the pin out, and jumped underneath the Iraqi tank, killing himself and disabling the tank. This stopped the Iraqi tank division's advance and saved many people's lives.	During the Iran and Iraq war, Hossein, a 13 year old boy, has sneaked into the army. A convoy of tanks confronts him. The tanks will destroy a part of the city where he is fighting. He has two options. The first option is, he can try to run to his commander to inform him about the situation. This would save his own life. The second option is, he can try to stop a tank. This would sacrifice his own life. Hossein took a grenade and pulled the pin. Then he jumped under a tank, killing himself and disabling the tank. This stopped the convoy's advance. Hossein saved many lives by choosing the second option.
Surface Δ	During the Bosnia and Serbian war, a young boy sneaks in to the army. One day during the war, he is confronted with a convoy of enemy buses carrying soldiers and weapons. If these buses are not stopped, they will help the enemy destroy part of the city that the boy is fighting at. He can either try to run to his commander on time, inform him about the situation and save his own life or he can stop a bus by running underneath it and activating a mine which otherwise would not work.	During the Bosnia and Serbian war, a young boy sneaks into the army. One day a convoy of buses that are carrying soldiers confronts him. The buses will help the enemy to destroy part of the city where he is fighting. He has two options. The first option is, he can try to run to his commander to inform him about the situation. This would save his own life. The second option is, he can run under a bus and activate a mine to stop the bus. This would sacrifice his own life.
Structural Δ	During a war, a young boy who has sneaked into the army, is confronted with a tank that if not stopped will destroy a part of the city that the boy is fighting at. He can either try to run to his commander on time and inform him about the attack which would cause the commander to issue a strike from other units against the tanks or he can stop one tank by sacrificing his own life.	During a war, a young boy, has sneaked into the army. A convoy of tanks confronts him. The tanks will destroy a part of the city where he is fighting. He has two options. The first option is, he can try to run to his commander to inform him about the situation. This would cause the commander to order another unit to attack the convoy. The second option is, hossein can try to stop a tank. This would sacrifice his own life.

Surface and Structural Δ	During the Bosnia and Serbian war, a young boy sneaks in to the army. One day during the war, he is confronted with an enemy bus carrying soldiers and weapons. If this bus is not stopped, it will help the enemy destroy part of the city that the boy is fighting at. He can either run to his commander on time, inform him about the situation which would cause the commander to issue a strike from other units against the convoy of buses or he can stop a bus by running underneath it and activating a mine which otherwise would not work.	During the Bosnia and Serbian war, a young boy sneaks into the army. One day a convoy of buses that are carrying soldiers confronts him. The buses will help the enemy to destroy part of the city where he is fighting. He has two options. The first option is, he can try to run to his commander to inform him about the situation. This would cause the commander to order another unit to attack the convoy. The second option is, the boy can run under a bus and activate a mine to stop the bus. This would sacrifice his own life.
Poryaie Vali		
Base	Pouryaie Vali was the most famous wrestler of his time. The morning before wrestling with a young athlete from another province, he goes to a mosque and sees the mother of the young athlete praying and saying "God, my son is going to wrestle with Pouryaie Vali. Please watch over him and help him win the match so he can use the prize money to buy a house". Pouryaie Vali thinks to himself that the young wrestler needs the money more than he does, and also winning the match will break the heart of the old mother. He has two choices, he can either win the match and keep his status as the best wrestler in the world or he could lose the match and make the old mother happy. Even though he was known not to ever lose a match, he loses that one on purpose.	Pouryaie Vali was the most famous wrestler of his time. He was going to wrestle a young athlete from another province. He goes to a mosque and sees the mother of the young athlete praying. She says, "My son is going to wrestle Pouryaie Vali. Please help him to win the match so that he can use the prize money to buy a house." Pouryaie Vali thinks that the young athlete needs the money more than he does. He also thinks that winning the match will break the old mother's heart. He has two options. The first option is, he can win the match. This would keep his status as the best wrestler. The second option is, he can lose the match to make the old mother happy. This would risk his status and help the young athlete to buy a house. He makes the old mother happy by choosing the second option.
Surface Δ	Ali is the greatest ping pong player of his city. The morning before a match with a young athlete from another city, he goes for a walk outside the stadium and sees the mother of the young athlete praying and saying "God, my son is going to play a match with Ali the famous ping pong player. Please watch over him and help him win the match so he can use the prize	Ali is the greatest ping-pong player of his city. He is going to play a young athlete from another city. He walks outside and sees the mother of the young athlete praying. She says, "My son is going to play a match with Ali. Please help him to win the match so that he can use the prize money to get married." Ali has two options. The first option is, he can win the match. This would

	money to get married". Ali has two choices, he can either win the match and keep his status as the best ping pong player or he could lose the match and make the old mother happy.	keep his status as the best ping pong player. The second option is, he can lose the match to make the old mother happy. This would risk his status and help the young athlete to get married.
Structural Δ	Ali was the most famous wrestler of his time. The morning before wrestling with a young athlete from another province, he goes to a mosque and sees the mother of the young athlete praying and saying "God, my son is going to wrestle with Ali. Please watch over him and help him win the match so he can use the prize to buy me new expensive clothes". Ali has two choices, he can either win the match and keep his status as the best wrestler in the world or he could lose the match and make the old mother happy.	Ali was the most famous wrestler of his time. He was going to wrestle a young athlete from another province. He goes to a mosque and sees the mother of the young athlete praying. She says, "My son is going to wrestle Ali. Please help him to win the match so that he can use the prize money to buy expensive clothes for me." Ali has two options. The first option is, he can win the match. This would keep his status as the best wrestler. The second option is, he can lose the match to make the old mother happy. This would risk his status and help the young athlete to buy expensive clothes for his mother.
Surface and Structural Δ	Ali is the greatest ping pong player of his city. The morning before a match with a young athlete from another city, he goes for a walk outside the stadium and sees the mother of the young athlete praying and saying "God, my son is going to play a match with Ali the famous ping pong player. Please watch over him and help him win the match so he can use the prize money use the prize to buy me new expensive clothes". Ali has two choices, he can either win the match and keep his status as the best ping pong player or he could lose the match and make the old mother happy.	Ali is the greatest ping-pong player of his city. He is going to play a young athlete from another city. He walks outside and sees the mother of the young athlete praying. She says, "My son is going to play a match with Ali. Please help him to win the match so that he can use the prize money to buy expensive clothes for me." Ali has two options. The first option is, he can win the match. This would keep his status as the best ping pong player. The second option is, he can lose the match to make the old mother happy. This would risk his status and help the young athlete to buy expensive clothes for his mother.

8.3 Appendix C: Fable interpretations

This section gives the full narrative functions and discourse-level DRS interpretations inferred for each fable dicussed in section 5.3.3. The DRS representations are displayed using indentation rather than box graphics.

8.3.1 The Dogs and the Fox

8.3.1.1 Narrative functions inferred by sentence

Sentence-3456749011-16493

```
(openExpectation
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17846)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  skin16567 group-of-dog16506 ?response))

(closedExpectation
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17845)
 (presentsResponse
  (PresentationEventFn Sentence-3456749014-17148 ?nent-meets) skin16567
  group-of-dog16506 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17845)
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17846)
 (presentsResponse
  (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17845)
  skin16567 group-of-dog16506 tear16760))

(introducesActor
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17847)
 lion16650)

(introducesActor
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17848)
 group-of-dog16506)

(presentsAwareness
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17846)
 group-of-dog16506 skin16567)

(presentsResponse
 (PresentationEventFn Sentence-3456749011-16493 IBTGeneration17845)
 skin16567 group-of-dog16506 tear16760)
```

Sentence-3456749014-17148

```
(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DrsCase17861)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DrsCase17861)) group-of-dog16506 ?action))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (subGoal (PresentationEventFn Sentence-3456749014-17148 ?nent-sets)
  group-of-dog16506 ?sub-goal ?goal))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase17861))))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase17861))))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DrsCase17861))))

(openExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17864)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  tear16760 fox17160 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17865)
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17865)
  find17428 (Goal-AvoidFn (DrsCaseFn DrsCase17861)) Failure))

(meetsExpectation
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17854)
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17864)
 (presentsResponse
  (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17854)
  tear16760 fox17160 say17241))

(introducesActor
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17866)
 fox17160)

(introducesGoal
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17863)
```

```
 group-of-dog16506 (Goal-AvoidFn (DrsCaseFn DrsCase17861)))

(evaluatesOutcome
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17865)
 find17428 (Goal-AvoidFn (DrsCaseFn DrsCase17861)) Failure)

(presentsAwareness
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17864)
 fox17160 tear16760)

(presentsResponse
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17854)
 tear16760 fox17160 say17241)

(presentsContrast
 (PresentationEventFn Sentence-3456749014-17148 IBTGeneration17867)
 (causes-Underspecified (hasExistentialStatus lion16650 Deceased)
  tear16760)
 (causes-Underspecified (hasExistentialStatus lion16650 Alive)
  find17428))
```

8.3.1.2 Discourse-level DRS

```
DRS-3456749017-17843

Universe: group-of-dog16506 begin16691 find16523 skin16567 lion16650 see17192 fox17160 say17241

(implies-DrsDrs (DrsCaseFn DRS-3456749035-17850)
 (DrsCaseFn DRS-3456749035-17851))


       DRS-3456749035-17850

       Universe: dog16506

       (member dog16506 group-of-dog16506)


       DRS-3456749035-17851

       (isa dog16506 Dog)


(temporallyIntersects find16523 begin16691)

(qualitativeExtent group-of-dog16506 Some)

(isa group-of-dog16506 Set-Mathematical)

(activityBegun begin16691 (DrsCaseFn DRS-3456749035-17852))


       DRS-3456749035-17852

       Universe: group-of-dog16506 tear16760 tooth16954

       (isa tear16760 Ripping)
```

```
        (objectOfStateChange tear16760 skin16567)

        (doneBy tear16760 group-of-dog16506)

        (possessiveRelation group-of-dog16506 tooth16954)

(isa skin16567 FurPelt)

(possessiveRelation lion16650 skin16567)

(anatomicalParts lion16650 skin16567)

(isa skin16567 CriticalOrgan)

(isa find16523 FindingSomething)

(objectFound find16523 skin16567)

(doneBy find16523 group-of-dog16506)

(isa begin16691 BeginningAnActivity)

(isa lion16650 Lion)

(performedBy begin16691 group-of-dog16506)

(detatched skin16567 lion16650)

(hasExistentialStatus lion16650 Deceased)

(isa lion16650 Agent-Generic)

(temporallyIntersects see17192 say17241)

(isa see17192 VisualPerception)

(performedBy see17192 fox17160)

(perceivedThings see17192 group-of-dog16506)

(isa fox17160 Fox)

(senderOfInfo say17241 fox17160)

(infoTransferred say17241 (DrsCaseFn DRS-3456749088-17868))

        DRS-3456749088-17868

        (implies-DrsDrs (DrsCaseFn DRS-3456749088-17869)
         (DrsCaseFn DRS-3456749088-17870))

               DRS-3456749088-17869

               Universe: be17368 lion16650

               (hasExistentialStatus lion16650 Alive)
```

```
(isa lion16650 Lion)

DRS-3456749088-17870

Universe: group-of-dog16506

(possible-Historical (DrsCaseFn DRS-3456749088-17871))

        DRS-3456749088-17871

        Universe: find17428

        (causes-SitProp find17428
         (knows group-of-dog16506 (DrsCaseFn DRS-3456749088-17872)))

               DRS-3456749088-17872

               Universe: group-of-dog16506 lion16650 claw17599 be17629 tooth17723

               (possessiveRelation group-of-dog16506 tooth17723)

               (possessiveRelation lion16650 claw17599)

               (isa claw17599 Claw)

               (isa tooth17723 Tooth)

               (isa be17629 ComparisonEvent)

               (comparer be17629 claw17599)

               (comparee be17629 tooth17723)

               (comparativeRelation be17629 Strong)

        (isa find17428 DiscoveringSomething)

        (nonDeliberateActors find17428 group-of-dog16506)

(recipientOfInfo say17241 group-of-dog16506)

(isa say17241 Informing)
```

8.3.2 The Boy and the Nettles

8.3.2.1 Narrative functions inferred by sentence

Sentence-3456759879-32846

```
(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DrsCase35325)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DrsCase35325)) boy32854 ?action))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) boy32854
  ?sub-goal ?goal))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  sting32882 ?actor ?response))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(openExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DrsCase35325)) ?outcome))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DrsCase35325)) boy32854 ?action))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) boy32854
  ?sub-goal (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
```

```
  (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(closedExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DrsCase35325))))

(meetsExpectation
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
  sting32882 (Goal-AvoidFn (DrsCaseFn DrsCase35325)) Failure))

(introducesActor
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35329)
 boy32854)

(introducesGoal
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35327)
 boy32854 (Goal-AvoidFn (DrsCaseFn DrsCase35325)))

(evaluatesOutcome
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 sting32882 (Goal-AvoidFn (DrsCaseFn DrsCase35325)) Failure)
```

Sentence-3456759880-33105

```
(openExpectation
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35333)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  tell33174 mother33231 ?response))

(closedExpectation
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35332)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  sting32882 boy32854 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35332)
 (PresentationEventFn Sentence-3456759879-32846 IBTGeneration35328)
 (presentsResponse
  (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35332)
  sting32882 boy32854 tell33174))

(introducesActor
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35335)
 mother33231)

(presentsAwareness
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35333)
 mother33231 tell33174)

(presentsResponse
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35332)
```

```
 sting32882 boy32854 tell33174)

(presentsContrast
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35337)
 (qualityOfAction touch33351 (MediumToHighAmountFn Gentleness))
 (qualityOfAction hurt33706 (HighToVeryHighAmountFn Intensity)))
```

Sentence-3456759887-34403

```
(openExpectation
 (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35343)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  reason34541 mother33231 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35342)
 (PresentationEventFn Sentence-3456759880-33105 IBTGeneration35333)
 (presentsResponse
  (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35342)
  tell33174 mother33231 say34418))

(presentsAwareness
 (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35343)
 mother33231 reason34541)

(presentsResponse
 (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35342)
 tell33174 mother33231 say34418)

(presentsContrast
 (PresentationEventFn Sentence-3456759887-34403 IBTGeneration35349)
 (causes-Underspecified touch33351 hurt33706)
 (causes-Underspecified grasp34934 (not hurt35075)))
```

8.3.2.2 Discourse-level DRS

```
DRS-3456759909-35324

Universe: be32863 sting32882 boy32854 nettle33021 run33114 home33147 tell33174 mother33231
say34418

(isa sting32882 Stinging)

(performedBy sting32882 nettle33021)

(objectActedOn sting32882 boy32854)

(isa boy32854 MaleChild)

(isa boy32854 Agent-Generic)

(possessiveRelation boy32854 home33147)
```

```
(possessiveRelation boy32854 mother33231)

(isa tell33174 Informing)

(infoTransferred tell33174 (DrsCaseFn DRS-3456759938-35340))

       DRS-3456759938-35340

       Universe: hurt33706 touch33351 boy32854 it33491 it33431

       (isa touch33351 TouchingEvent)

       (isa hurt33706 HarmingAnAgent)

       (maleficiary hurt33706 boy32854)

       (performedBy touch33351 boy32854)

       (qualityOfAction touch33351 (MediumToHighAmountFn Gentleness))

       (qualityOfAction hurt33706 (HighToVeryHighAmountFn Intensity))

       (doneBy hurt33706 it33491)

(isa mother33231 HumanMother)

(senderOfInfo tell33174 boy32854)

(recipientOfInfo tell33174 mother33231)

(isa mother33231 Agent-Generic)

(isa say34418 Informing)

(senderOfInfo say34418 mother33231)

(infoTransferred say34418
 (and (DrsCaseFn DRS-3456759967-35353)
      (DrsCaseFn DRS-3456759967-35354)))

       DRS-3456759967-35353

       Universe: tell33174

       (causes-Underspecified that34499 (DrsCaseFn DRS-3456759967-35355))

              DRS-3456759967-35355

              Universe: it34600 boy32854 sting34611

              (isa sting34611 StingingByAnimal)

       DRS-3456759967-35354
```

```
Universe: boy32854 whenever34795

(ruleQualifier whenever34795 (DrsCaseFn DRS-3456759967-35356))

        DRS-3456759967-35356

        Universe: it35001 it34975 grasp34934

        (willBe (DrsCaseFn DRS-3456759967-35357))

                DRS-3456759967-35357

                (not (DrsCaseFn DRS-3456759967-35358))

                        DRS-3456759967-35358

                        Universe: hurt35075 boy32854

                        (isa hurt35075 HarmingAnAgent)

                        (maleficiary hurt35075 boy32854)

                        (doneBy hurt35075 it35001)

                (isa grasp34934 GraspingSomething)

        (isa whenever34795 TouchingEvent)

(recipientOfInfo say34418 boy32854)
```

8.3.3 The Boys and the Frogs

8.3.3.1 Narrative functions inferred by sentence

Sentence-3456965386-23233

```
(openExpectation
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26608)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  group-of-frog23619 group-of-boy23246 ?response))

(closedExpectation
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26611)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  group-of-frog23619 group-of-boy23246 ?response))

(meetsExpectation
```

```
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26611)
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26608)
 (presentsResponse
  (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26611)
  group-of-frog23619 group-of-boy23246 pelt24195))

(introducesActor
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26613)
 group-of-frog23619)

(introducesActor
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26614)
 group-of-boy23246)

(establishesActivity
 (PresentationEventFn Sentence-3456965386-23233 ?narrative-event)
 group-of-boy23246 play23255)

(presentsAwareness
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26608)
 group-of-boy23246 group-of-frog23619)

(presentsResponse
 (PresentationEventFn Sentence-3456965386-23233 IBTGeneration26611)
 group-of-frog23619 group-of-boy23246 pelt24195)
```

Sentence-3456965397-25826

```
(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DrsCase26624)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DrsCase26624)) group-of-frog23619 ?action))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets)
  group-of-frog23619 ?sub-goal ?goal))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26627)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  kill25837 ?actor ?response))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase26624))))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
```

```
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DrsCase26624))))

(openExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DrsCase26624))))

(meetsExpectation
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26627)
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26627)
  kill25837 (Goal-AvoidFn (DrsCaseFn DrsCase26624)) Failure-Partial))

(introducesGoal
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 group-of-frog23619 (Goal-AvoidFn (DrsCaseFn DrsCase26624)))

(evaluatesOutcome
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26627)
 kill25837 (Goal-AvoidFn (DrsCaseFn DrsCase26624)) Failure-Partial)

(presentsResult
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26628)
 begin23949 kill25837)
```

Sentence-3456965398-25923

```
(openExpectation
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26633)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  say26187 group-of-boy23246 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26640)
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26626)
 (presentsAction
  (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26640)
  (Goal-AvoidFn (DrsCaseFn DrsCase26624)) group-of-frog23619 say26187))

(meetsExpectation
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26641)
 (PresentationEventFn Sentence-3456965397-25826 IBTGeneration26627)
 (presentsResponse
  (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26641)
  kill25837 group-of-frog23619 say26187))

(introducesActor
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26634)
 head26003)

(presentsAction
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26640)
 (Goal-AvoidFn (DrsCaseFn DrsCase26624)) group-of-frog23619 say26187)
```

```
(presentsAwareness
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26633)
 group-of-boy23246 say26187)

(presentsResponse
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26641)
 kill25837 group-of-frog23619 say26187)

(presentsContrast
 (PresentationEventFn Sentence-3456965398-25923 IBTGeneration26639)
 (perspectives group-of-boy23246 (isa death26528 Sport))
 (perspectives group-of-frog23619 (isa death26528 Dying)))
```

8.3.3.2 Discourse-level DRS

```
DRS-3456965409-26607

Universe: begin23949 see23440 pond23382 group-of-boy23246 play23255 group-of-frog23619 subset-of-
frog25885 kill25837 water26105 one-of-frog25959 lift25970 say26187 head26003 his25992

(temporallyIntersects play23255 (cconj see23440 begin23949))

(isa group-of-boy23246 Set-Mathematical)

(implies-DrsDrs (DrsCaseFn DRS-3456965426-26615)
 (DrsCaseFn DRS-3456965426-26616))


       DRS-3456965426-26615

       Universe: boy23246

       (member boy23246 group-of-boy23246)


       DRS-3456965426-26616

       (isa boy23246 MaleChild)


(qualitativeExtent group-of-boy23246 Some)

(isa see23440 VisualPerception)

(isa begin23949 BeginningAnActivity)

(isa group-of-frog23619 Set-Mathematical)

(isa play23255 RecreationalActivity)

(performedBy see23440 group-of-boy23246)

(perceivedThings see23440 group-of-frog23619)

(activityBegun begin23949 (DrsCaseFn DRS-3456965426-26617))
```

```
        DRS-3456965426-26617

        Universe: group-of-frog23619 pelt24195 group-of-stone25092

        (isa pelt24195 ThrowingAndHittingSomething)

        (isa group-of-stone25092 Set-Mathematical)

        (performedBy pelt24195 group-of-boy23246)

        (target pelt24195 group-of-frog23619)

        (implies-DrsDrs (DrsCaseFn DRS-3456965426-26618)
         (DrsCaseFn DRS-3456965426-26619))

                DRS-3456965426-26618

                Universe: stone25092

                (member stone25092 group-of-stone25092)

                DRS-3456965426-26619

                (transferredThing pelt24195 group-of-stone25092)

        (implies-DrsDrs (DrsCaseFn DRS-3456965426-26620)
         (DrsCaseFn DRS-3456965426-26621))

                DRS-3456965426-26620

                Universe: frog23619

                (member frog23619 group-of-frog23619)

                DRS-3456965426-26621

                (isa frog23619 Frog)

(performedBy play23255 group-of-boy23246)

(performedBy begin23949 group-of-boy23246)

(qualitativeExtent group-of-frog23619 Some)

(subsetOf subset-of-frog25885 group-of-frog23619)

(implies-DrsDrs (DrsCaseFn DRS-3456965451-26629)
 (DrsCaseFn DRS-3456965451-26630))

        DRS-3456965451-26629

        Universe: elt-of-frog25885
```

```
        (member elt-of-frog25885 group-of-frog23619)

        DRS-3456965451-26630

        Universe: elt-of-frog25885

        (isa elt-of-frog25885 Frog)

(isa subset-of-frog25885 Set-Mathematical)

(qualitativeExtent subset-of-frog25885 Some)

(isa kill25837 KillingByOrganism)

(organismKilled kill25837 subset-of-frog25885)

(performedBy kill25837 group-of-boy23246)

(member one-of-frog25959 group-of-frog23619)

(temporallyIntersects lift25970 say26187)

(implies-DrsDrs (DrsCaseFn DRS-3456967760-26642)
 (DrsCaseFn DRS-3456967760-26643))

        DRS-3456967760-26642

        Universe: elt-of-frog25959

        (member elt-of-frog25959 group-of-frog23619)

        DRS-3456967760-26643

        Universe: elt-of-frog25959

        (isa elt-of-frog25959 Frog)

(possessiveRelation his25992 head26003)

(infoTransferred say26187
 (and (DrsCaseFn DRS-3456967760-26644)
      (DrsCaseFn DRS-3456967760-26645)))

        DRS-3456967760-26644

        Universe: stop26449

        (isa stop26449 DiscontinuingAnActivity)

        (interrupts stop26449 (GAP SUBJECT))
```

```
DRS-3456967760-26645

Universe: sport26489 group-of-boy23246 group-of-frog23619

(isa sport26489 Dying)

(possessiveRelation group-of-boy23246 sport26489)

(isa sport26489 Sport)

(bodilyDoer sport26489 group-of-frog23619)
```

```
(isa head26003 Leader)

(isa say26187 Informing)

(senderOfInfo say26187 one-of-frog25959)

(recipientOfInfo say26187 group-of-boy23246)
```

8.3.4 The Cat and Venus

8.3.4.1 Narrative functions inferred by sentence

Sentence-3457351679-52592

```
(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57075)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57074)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57075)) cat52600 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57074)) cat52600 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) cat52600
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57075))))

(openExpectation
```

```
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) cat52600
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57075))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57075))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57075))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57081)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  (requests cat52600 Venus-TheGoddess
   (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))
  Venus-TheGoddess ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57082)
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (presentsAction
  (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57082)
  (Goal-AchieveFn (DrsCaseFn DRS57074)) cat52600
  (requests cat52600 Venus-TheGoddess
   (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))))

(meetsExpectation
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (subGoal
  (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
  cat52600 (Goal-AchieveFn (DrsCaseFn DRS57075))
  (Goal-AchieveFn (DrsCaseFn DRS57074))))

(introducesActor
```

```
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57083)
 cat52600)

(introducesActor
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57084)
 man52707)

(introducesActor
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57085)
 Venus-TheGoddess)

(introducesGoal
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 cat52600 (Goal-AchieveFn (DrsCaseFn DRS57075)))

(introducesGoal
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 cat52600 (Goal-AchieveFn (DrsCaseFn DRS57074)))

(presentsAction
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57082)
 (Goal-AchieveFn (DrsCaseFn DRS57074)) cat52600
 (requests cat52600 Venus-TheGoddess
  (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess)))

(subGoal
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 cat52600 (Goal-AchieveFn (DrsCaseFn DRS57075))
 (Goal-AchieveFn (DrsCaseFn DRS57074)))

(presentsAwareness
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57081)
 Venus-TheGoddess
 (requests cat52600 Venus-TheGoddess
  (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess)))
```

Sentence-3457351682-53141

```
(openExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  change53365 ?actor ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57075)) ?outcome))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57075)) cat52600 ?action))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
```

```
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) cat52600
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57075))))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57075))))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57075))))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57075))))

(closedExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57090)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  (requests cat52600 Venus-TheGoddess
   (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))
  Venus-TheGoddess ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57076)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
  change53365 (Goal-AchieveFn (DrsCaseFn DRS57075)) Success))

(meetsExpectation
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57090)
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57081)
 (presentsResponse
  (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57090)
  (requests cat52600 Venus-TheGoddess
   (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))
  Venus-TheGoddess change53365))

(introducesActor
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57091)
 woman53542)

(evaluatesOutcome
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57089)
 change53365 (Goal-AchieveFn (DrsCaseFn DRS57075)) Success)

(presentsResult
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57092)
 consent53152 change53365)

(presentsResponse
 (PresentationEventFn Sentence-3457351682-53141 IBTGeneration57090)
 (requests cat52600 Venus-TheGoddess
  (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))
 Venus-TheGoddess change53365)
```

Sentence-3457351683-53641

```
(openExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  marry54128 ?actor ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57074)) ?outcome))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57074)) cat52600 ?action))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) cat52600
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57074))))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57074))))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57074))))

(closedExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57074))))

(meetsExpectation
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 (PresentationEventFn Sentence-3457351679-52592 IBTGeneration57078)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
  marry54128 (Goal-AchieveFn (DrsCaseFn DRS57074)) Success))

(evaluatesOutcome
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57100)
 marry54128 (Goal-AchieveFn (DrsCaseFn DRS57074)) Success)

(presentsResult
 (PresentationEventFn Sentence-3457351683-53641 IBTGeneration57101)
 love53863 marry54128)
```

Sentence-3457351685-54407

```
(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57107)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57107)) Venus-TheGoddess ?action))

(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) Venus-TheGoddess
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57107))))

(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57107))))

(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57107))))

(openExpectation
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57107))))

(introducesGoal
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 Venus-TheGoddess (Goal-AchieveFn (DrsCaseFn DRS57107)))

(presentsContrast
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57116)
 (objectOfStateChange change54572 shape54745)
 (objectOfStateChange change54951 habit55136))
```

Sentence-3457351688-55380

```
(openExpectation
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57123)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  mouse55755 Couple-PairOfPeople57077 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57120)
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (presentsAction
```

```
  (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57120)
  (Goal-AchieveFn (DrsCaseFn DRS57107)) Venus-TheGoddess place55685))

(introducesActor
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57124)
 mouse55755)

(presentsAction
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57120)
 (Goal-AchieveFn (DrsCaseFn DRS57107)) Venus-TheGoddess place55685)

(presentsAwareness
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57123)
 Couple-PairOfPeople57077 mouse55755)
```

Sentence-3457351696-56135

```
(openExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  pursue56303 ?actor ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS57107)) ?outcome))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS57107)) Venus-TheGoddess ?action))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) Venus-TheGoddess
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS57107))))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57107))))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS57107))))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS57107))))

(closedExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57130)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
```

```
  mouse55755 Couple-PairOfPeople57077 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (PresentationEventFn Sentence-3457351685-54407 IBTGeneration57108)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
  pursue56303 (Goal-AchieveFn (DrsCaseFn DRS57107)) Failure))

(meetsExpectation
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57130)
 (PresentationEventFn Sentence-3457351688-55380 IBTGeneration57123)
 (presentsResponse
  (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57130)
  mouse55755 Couple-PairOfPeople57077 pursue56303))

(evaluatesOutcome
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 pursue56303 (Goal-AchieveFn (DrsCaseFn DRS57107)) Failure)

(presentsResult
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57131)
 place55685 pursue56303)

(presentsResult
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57132)
 forget56154 pursue56303)

(presentsResponse
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57130)
 mouse55755 Couple-PairOfPeople57077 pursue56303)
```

Sentence-3457351698-56591

```
(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-MaintainFn (DrsCaseFn DRS57074)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-MaintainFn (DrsCaseFn DRS57074)) cat52600 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) cat52600
  ?sub-goal (Goal-MaintainFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57150)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  return56744 ?actor ?response))

(openExpectation
```

```
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-MaintainFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-MaintainFn (DrsCaseFn DRS57074))))

(openExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-MaintainFn (DrsCaseFn DRS57074))))

(meetsExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57145)
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (presentsResponse
  (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57145)
  pursue56303 Venus-TheGoddess return56744))

(meetsExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57143)
 (PresentationEventFn Sentence-3457351696-56135 IBTGeneration57128)
 (presentsResponse
  (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57143)
  pursue56303 Venus-TheGoddess
  (feelsEmotion Venus-TheGoddess
   (MediumToVeryHighAmountFn Disappointment))))

(meetsExpectation
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57150)
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57150)
  return56744 (Goal-MaintainFn (DrsCaseFn DRS57074)) Failure))

(introducesGoal
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57149)
 cat52600 (Goal-MaintainFn (DrsCaseFn DRS57074)))

(evaluatesOutcome
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57150)
 return56744 (Goal-MaintainFn (DrsCaseFn DRS57074)) Failure)

(presentsResult
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57151)
 be56601 return56744)

(presentsResponse
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57143)
 pursue56303 Venus-TheGoddess
 (feelsEmotion Venus-TheGoddess
  (MediumToVeryHighAmountFn Disappointment)))

(presentsResponse
 (PresentationEventFn Sentence-3457351698-56591 IBTGeneration57145)
 pursue56303 Venus-TheGoddess return56744)
```

8.3.4.2 Discourse-level DRS

```
DRS-3457351709-57072

Universe: ask52756 man52707 cat52600 love52610 woman53542 change53365 consent53152 cause53665
wish54418 Couple-PairOfPeople57077 chamber55600 their55595 middle55843 place55685 Venus-
TheGoddess recline55483 mouse55755 pursue56303 her56178 forget56154 condition56257 her56908
return56744 be56601 shape56971 series56954 current56955

(loves cat52600 man52707)

(requests cat52600 Venus-TheGoddess
 (doneBy (DrsCaseFn DRS57075) Venus-TheGoddess))


       DRS57075

       Universe: cat52600 change52835 woman53037

       (performedBy change52835 Venus-TheGoddess)

       (objectOfStateChange change52835 cat52600)

       (isa change52835 IntrinsicStateChangeEvent)


(isa cat52600 Agent-Generic)

(isa cat52600 Cat)

(isa man52707 Agent-Generic)

(isa man52707 AdultMaleHuman)

(isa woman53542 AdultFemaleHuman)

(isa change53365 IntrinsicStateChangeEvent)

(isa consent53152 MakingAnAgreement)

(objectOfStateChange change53365 cat52600)

(into-UnderspecifiedContainer change53365 woman53542)

(performedBy change53365 Venus-TheGoddess)

(requestStatement consent53152 (DrsCaseFn DRS-3457351762-57097))


       DRS-3457351762-57097

       Universe: her53240 request53288 fulfill53199

       (possessiveRelation her53240 request53288)


(agreeingAgents consent53152 Venus-TheGoddess)

(isa woman53542 Agent-Generic)
```

```
(causes-SitProp change53365 (DrsCaseFn DRS-3457351775-57106))

        DRS-3457351775-57106

        Universe: cat52600 man52707 marry54128 her54029 love53863

        (isa marry54128 WeddingEvent-Entire)

        (eventHonors marry54128 man52707)

        (eventHonors marry54128 cat52600)

(desires Venus-TheGoddess (DrsCaseFn DRS57107))

        DRS57107

        Universe: discover54471

        (conditionEvaluated discover54471 (DrsCaseFn DRS-3457351805-57117))

                DRS-3457351805-57117

                Universe: cat52600 shape54745 life55227 habit55136 change54951 change54572

                (isa change54951 IntrinsicStateChangeEvent)

                (isa change54572 IntrinsicStateChangeEvent)

                (performedBy change54572 cat52600)

                (performedBy change54951 cat52600)

                (objectOfStateChange change54572 shape54745)

                (objectOfStateChange change54951 habit55136)

                (possessiveRelation cat52600 shape54745)

                (possessiveRelation cat52600 habit55136)

                (after change54951 change54572)

                (isa habit55136 RoutineBehavior)

        (isa discover54471 CheckingWhetherConditionObtains)

        (performedBy discover54471 Venus-TheGoddess)

(possessiveRelation their55595 chamber55600)

(startsDuring place55685 recline55483)
```

```
(isa recline55483 Situation)

(isa place55685 CausingAnotherObjectsTranslationalMotion)

(isa Couple-PairOfPeople57077 Couple-PairOfPeople)

(isa chamber55600 RoomInAConstruction)

(isa mouse55755 Mouse-Rodent)

(isa chamber55600 Bedroom)

(providerOfMotiveForce place55685 Venus-TheGoddess)

(primaryObjectMoving place55685 mouse55755)

(toLocation place55685 middle55843)

(centerOf middle55843 chamber55600)

(holdsIn recline55483
 (objectFoundInLocation Couple-PairOfPeople57077 chamber55600))

(possessiveRelation chamber55600 middle55843)

(holdsIn recline55483 (isa Couple-PairOfPeople57077 RecliningPosture))

(isa mouse55755 Agent-Generic)

(possessiveRelation her56178 condition56257)

(purposeInEvent cat52600 pursue56303 (DrsCaseFn DRS-3457351975-57142))


       DRS-3457351975-57142

       Universe: eat56417 it56480

       (isa eat56417 EatingEvent)


(temporallyIntersects forget56154 pursue56303)

(isa pursue56303 PursuingSomething)

(performedBy pursue56303 cat52600)

(objectPursued pursue56303 mouse55755)

(possessiveRelation her56908 shape56971)

(isa return56744 IntrinsicStateChangeEvent)

(isa shape56971 PhysicalForm-Underspecified)

(feelsEmotion Venus-TheGoddess
 (MediumToVeryHighAmountFn Disappointment))

(performedBy return56744 Venus-TheGoddess)
```

```
(objectOfStateChange return56744 cat52600)

(toState return56744 shape56971)

(priorInSeries shape56971 current56955 series56954)

(occursDuring current56955 Now)
```

8.3.5 The Dove and the Ant

8.3.5.1 Narrative functions inferred by sentence

Sentence-3457354751-66199

```
(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AchieveFn (DrsCaseFn DRS68444)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AchieveFn (DrsCaseFn DRS68444)) ant66204 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) ant66204
  ?sub-goal (Goal-AchieveFn (DrsCaseFn DRS68444))))

(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS68444))))

(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AchieveFn (DrsCaseFn DRS68444))))

(openExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AchieveFn (DrsCaseFn DRS68444))))

(meetsExpectation
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68446)
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 (presentsAction
  (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68446)
  (Goal-AchieveFn (DrsCaseFn DRS68444)) ant66204 go66210))

(introducesActor
```

```
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68447)
 ant66204)

(introducesGoal
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68445)
 ant66204 (Goal-AchieveFn (DrsCaseFn DRS68444)))

(presentsAction
 (PresentationEventFn Sentence-3457354751-66199 IBTGeneration68446)
 (Goal-AchieveFn (DrsCaseFn DRS68444)) ant66204 go66210)
```

Sentence-3457354752-66372

```
(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DRS68451)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DRS68451)) ant66204 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) ant66204
  ?sub-goal (Goal-AvoidFn (DrsCaseFn DRS68451))))

(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68451))))

(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68451))))

(openExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DRS68451))))

(meetsExpectation
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68453)
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (introducesThreat
  (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68453)
  carry66484 (Goal-AvoidFn (DrsCaseFn DRS68451))))

(introducesGoal
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 ant66204 (Goal-AvoidFn (DrsCaseFn DRS68451)))

(introducesThreat
```

```
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68453)
 carry66484 (Goal-AvoidFn (DrsCaseFn DRS68451)))
```

Sentence-3457354754-66723

```
(openExpectation
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68460)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  carry66484 dove66731 ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68461)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  carry66484 dove66731 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68459)
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (introducesOpportunity
  (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68459)
  (in-Generic bough66870 river66249)
  (Goal-AvoidFn (DrsCaseFn DRS68451))))

(meetsExpectation
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68461)
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68460)
 (presentsResponse
  (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68461)
  carry66484 dove66731 throw66807))

(introducesActor
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68462)
 dove66731)

(introducesOpportunity
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68459)
 (in-Generic bough66870 river66249) (Goal-AvoidFn (DrsCaseFn DRS68451)))

(presentsAwareness
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68460)
 dove66731 carry66484)

(presentsResponse
 (PresentationEventFn Sentence-3457354754-66723 IBTGeneration68461)
 carry66484 dove66731 throw66807)
```

Sentence-3457354756-66996

```
(openExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
```

```
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  reach67109 ?actor ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DRS68451)) ?outcome))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DRS68451)) ant66204 ?action))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) ant66204
  ?sub-goal (Goal-AvoidFn (DrsCaseFn DRS68451))))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68451))))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68451))))

(closedExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DRS68451))))

(meetsExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
  reach67109 (Goal-AvoidFn (DrsCaseFn DRS68451)) Success))

(meetsExpectation
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68466)
 (PresentationEventFn Sentence-3457354752-66372 IBTGeneration68452)
 (presentsAction
  (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68466)
  (Goal-AvoidFn (DrsCaseFn DRS68451)) ant66204 use67017))

(presentsAction
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68466)
 (Goal-AvoidFn (DrsCaseFn DRS68451)) ant66204 use67017)

(evaluatesOutcome
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68464)
 reach67109 (Goal-AvoidFn (DrsCaseFn DRS68451)) Success)

(presentsResult
 (PresentationEventFn Sentence-3457354756-66996 IBTGeneration68467)
 use67017 reach67109)
```

Sentence-3457354757-67256

```
(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (evaluatesOutcome (PresentationEventFn ?sid-meets ?nent-meets) ?event
  (Goal-AvoidFn (DrsCaseFn DRS68472)) ?outcome))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (presentsAction (PresentationEventFn ?sid-meets ?nent-meets)
  (Goal-AvoidFn (DrsCaseFn DRS68472)) dove66731 ?action))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (subGoal (PresentationEventFn ?sid-meets ?nent-meets) dove66731
  ?sub-goal (Goal-AvoidFn (DrsCaseFn DRS68472))))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (introducesThreat (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68472))))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (introducesObstacle (PresentationEventFn ?sid-meets ?nent-meets) ?sit
  (Goal-AvoidFn (DrsCaseFn DRS68472))))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (introducesOpportunity (PresentationEventFn ?sid-meets ?nent-meets)
  ?sit (Goal-AvoidFn (DrsCaseFn DRS68472))))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68476)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  aim67405 ant66204 ?response))

(openExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68475)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  man67348 ant66204 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68477)
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (introducesThreat
  (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68477)
  aim67405 (Goal-AvoidFn (DrsCaseFn DRS68472))))

(introducesActor
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68478)
 man67348)

(introducesGoal
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
```

```
 dove66731 (Goal-AvoidFn (DrsCaseFn DRS68472)))

(introducesThreat
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68477)
 aim67405 (Goal-AvoidFn (DrsCaseFn DRS68472)))

(presentsAwareness
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68475)
 ant66204 man67348)

(presentsAwareness
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68476)
 ant66204 aim67405)

(presentsSymmetry
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68481)
 (threatens carry66484 ant66204) (threatens aim67405 dove66731))
```

Sentence-3457354761-68013

```
(openExpectation
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68485)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  miss68188 ?actor ?response))

(closedExpectation
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68489)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  aim67405 ant66204 ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68485)
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68485)
  miss68188 (Goal-AvoidFn (DrsCaseFn DRS68472)) Success-Partial))

(meetsExpectation
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68489)
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68476)
 (presentsResponse
  (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68489)
  aim67405 ant66204 sting68034))

(evaluatesOutcome
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68485)
 miss68188 (Goal-AvoidFn (DrsCaseFn DRS68472)) Success-Partial)

(presentsResult
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68491)
 sting68034 miss68188)

(presentsResponse
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68489)
 aim67405 ant66204 sting68034)
```

```
(presentsSymmetry
 (PresentationEventFn Sentence-3457354761-68013 IBTGeneration68490)
 (responds carry66484 dove66731 throw66807)
 (response aim67405 ant66204 sting68034))
```

Sentence-3457354762-68277

```
(openExpectation
 (PresentationEventFn Sentence-3457354762-68277 IBTGeneration68494)
 (presentsResponse (PresentationEventFn ?sid-meets ?nent-meets)
  sting68034 ?actor ?response))

(meetsExpectation
 (PresentationEventFn Sentence-3457354762-68277 IBTGeneration68494)
 (PresentationEventFn Sentence-3457354757-67256 IBTGeneration68473)
 (evaluatesOutcome
  (PresentationEventFn Sentence-3457354762-68277 IBTGeneration68494)
  sting68034 (Goal-AvoidFn (DrsCaseFn DRS68472)) Success))

(evaluatesOutcome
 (PresentationEventFn Sentence-3457354762-68277 IBTGeneration68494)
 sting68034 (Goal-AvoidFn (DrsCaseFn DRS68472)) Success)

(presentsResult
 (PresentationEventFn Sentence-3457354762-68277 IBTGeneration68495)
 see67293 sting68034)
```

8.3.5.2 Discourse-level DRS

```
DRS-3457354776-68443

Universe: ant66204 go66210 river66249 drink66277 be66459 fall66381 carry66484 dove66731
throw66807 bough66870 pity66739 use67017 shore67175 reach67109 see67293 man67348 gun67492
aim67405 cause68137 foot68103 sting68034 miss68188 life68398 save68298

(purposeInEvent ant66204 go66210 (DrsCaseFn DRS68444))

       DRS68444

       Universe: drink66277

       (isa drink66277 DrinkingEvent)

       (performedBy drink66277 ant66204)

(isa ant66204 Ant)

(isa go66210 Movement-TranslationEvent)

(primaryObjectMoving go66210 ant66204)
```

```
(isa ant66204 Agent-Generic)

(isa river66249 FlowPath)

(isa river66249 River)

(isa fall66381 FallingEvent)

(missingSemTrans along66525 advpart Along-TheWord)

(isa carry66484 Conveying-Stationary)

(isa drink66277 FluidFlow-Translation)

(conveyor-Stationary carry66484 drink66277)

(into-UnderspecifiedContainer fall66381 river66249)

(isa fall66381 Accident)

(properParts drink66277 river66249)

(enables fall66381 carry66484)

(unintendedSituation carry66484)

(isa carry66484 DangerousSituation)

(transportees carry66484 ant66204)

(threaten-TowardsProp ant66204 carry66484 (DrsCaseFn DRS68451))

       DRS68451

       (isa Drowning68449 Drowning)

       (bodilyDoer Drowning68449 ant66204)

(bodilyActedOn carry66484 ant66204)

(eventOccursAt carry66484 river66249)

(possessiveRelation ant66204 carry66484)

(isa throw66807 ThrowingAnObject)

(isa pity66739 Situation)

(isa dove66731 Dove)

(isa carry66484 ExperiencingSomething)

(holdsIn pity66739
 (feelsTowardsEvent dove66731 carry66484 Pity positiveAmountOf))

(bodilyDoer carry66484 ant66204)
```

```
(performedBy throw66807 dove66731)

(objectActedOn throw66807 bough66870)

(into-UnderspecifiedContainer throw66807 river66249)

(relocateInstrument throw66807 bough66870 river66249)

(purposeInEvent ant66204 use67017 reach67109)

(isa bough66870 TreeBranch)

(purposeInEvent ant66204 use67017 (DrsCaseFn reach67109))

(isa reach67109 ArrivingAtAPlace)

(primaryObjectMoving reach67109 ant66204)

(toLocation reach67109 shore67175)

(isa use67017 UsingAnObject)

(instrument-Generic use67017 bough66870)

(performedBy use67017 ant66204)

(isa see67293 VisualPerception)

(isa man67348 AdultMaleHuman)

(isa gun67492 Gun)

(isa aim67405 AimingSomething)

(performedBy aim67405 man67348)

(objectActedOn aim67405 gun67492)

(at-UnderspecifiedLandmark aim67405 dove66731)

(threaten-TowardsProp dove66731 aim67405 (DrsCaseFn DRS68472))

       DRS68472

       (isa ShootingAndHittingSomething68469 ShootingAndHittingSomething)

       (performedBy ShootingAndHittingSomething68469 man67348)

       (damages ShootingAndHittingSomething68469 dove66731)

       (causes-SitSit ShootingAndHittingSomething68469 Dying68470)

       (isa Dying68470 Dying)

       (bodilyDoer Dying68470 dove66731)

(performedBy see67293 ant66204)
```

```
(perceivedThings see67293 man67348)

(causes-SitSit sting68034 miss68188)

(isa man67348 Agent-Generic)

(isa sting68034 StingingByAnimal)

(isa miss68188 (AttemptingFn HittingAnObject))

(performedBy miss68188 man67348)

(failureForAgents miss68188 man67348)

(causes-SitProp sting68034 (DrsCaseFn miss68188))

(performedBy sting68034 ant66204)

(objectActedOn sting68034 man67348)

(in-UnderspecifiedContainer sting68034 foot68103)

(possessiveRelation dove66731 life68398)

(isa life68398 Living)

(isa save68298 RescuingSomeone)

(performedBy save68298 sting68034)

(beneficiary save68298 life68398)
```