What can linguistics contribute to event extraction?

Charles J. Fillmore

Srini Narayanan

Collin F. Baker
International Computer Science Institute
Berkeley, California
{fillmore,snarayan,collinb}@icsi.berkeley.edu

Abstract

This paper examines the question of how a linguistic analysis of a written document can contribute to identifying, tracking and populating the "eventualities" that are presented in the document, either directly or indirectly, and representing degrees of belief concerning them. It is our view that the role of lexical analysis (as exemplified in the research carried out in the FrameNet project) is greater than usually assumed, so this paper is partly an attempt to clarify the boundary between on the one hand the information that can be derived on the basis of linguistic knowledge alone (composed of lexical meanings and the meanings of grammatical constructions) and on the other hand, reasoning based on beliefs about the source of a document, world knowledge, and "common sense".

Since the general linguistic processes described in this paper will apply to eventualities in general (by which we mean acts, happenings, states of affairs, and relations, whether real, proposed, imagined, or denied), our presentation will emphasize the linguistic processes themselves. In particular, we show that the kind of information produced by the lexicon-building project FrameNet can have a special role in contributing to text understanding, starting from the basic facts of the combinatorial properties of frame-bearing words (verbs, nouns, adjectives and prepositions) and arriving at the means of recognizing the anaphoric properties of specific **unexpressed** event participants, for all parts of speech, in defining a new layer of anaphora resolution and text cohesion. Using as a starting point the challenge text presented in the call for this workshop (hereafter referred to as the Hijacking text), we show the points at which a thorough linguistic analysis can articulate with the kind of simulation formalism demonstrated in X-schema diagram, Figure 2 , which itself incorporates a great deal of world knowledge connected with the events introduced in the Hijacking text.

Valence and Text Understanding

We will begin our discussion of the ongoing FrameNet lexicon-building work with those parts that more or less match familiar linguistic analysis and are suffi cient for dealing with the linguistic analysis of the Hijacking text, and then survey further kinds of linguistic information that (1) make it possible to recognize attributions of truth claims, or (2) identify entities and eventualities that are not explicitly mentioned in given sentences.

The FrameNet project (`http://framenet.icsi.berkeley.edu`, Fontenelle 2003) is devoted to discovering and describing the lexical valences of lexical units in English, that is, their semantic and syntactic combinatorial properties, and how these properties can be used for identifying and populating the eventualities that are linguistically coded in a document. The most straightforward way in which eventualities can be found and fi lled in is through (1) the recognition of frame-bearing words that designate eventualities of particular types and (2) the identifi cation of phrases in the syntactic context of such words that denote participants ("slot fi llers") in these eventualities.

Consider the sentence *The bodyguard poisoned the emperor*; here, the transitive verb *poison* designates an event type in which a person administers poison to a living being, and this sentence identifi es the person referred to as *the bodyguard* as the agent of such an act, and the person referred to as *the emperor* as its victim. Given that the sentence has a past-tense form, we can say that the document claims this to be an actual occurrence, and the sentence invites the expectation (through the use of *the* and the simple past tense) that an earlier part of the text offers further details about the individuals and the time of the event.

The adjective *fond*, as in *Our daughters are fond of horses*, designates a particular positive psychological state, and in this sentence the holder of the attitude is identifi ed as our daughters, and the object of this state is taken to be horses in general. As a present-tense sentence it can be seen as claiming that this is a more or less permanent disposition of the individuals referred to as our daughters.

The noun *advice*, found in *John's advice to his daughter*, designates a conversational event of advice-giving, and in this sentence John is presented as the communicator in an instance of this event type and his daughter is the addressee. Since this is a noun phrase, temporal coordinates are missing, but in a past-tense referring context like *John's advice to his daughter was wise*, we can say that the text claims that such an event did occurr. In each case further details about the events and the participants, including the content of the advice, are likely to be found in surrounding parts of the text.

Almost all of the complexity of the Hijacking text can be dealt with through the interpretation of its frame-bearing verbs and nouns; the question of event order requires fur-

ther kinds of reasoning and evidence, some (but not all) of it based in the sentence's linguistic form. In the Hijacking sentence,

(1) The United Nations says Somali gunmen who hijacked a U.N.-chartered vessel carrying food aid for tsunami victims have released the ship after holding it for more than two months.

the main frame-bearing lexical units that evoke event types are *says, hijacked, chartered, carrying, aid, tsunami, victims, released*, and *holding*. Named entities in the sentence include *United Nations*, *Somali*, and *U.N.*; names of other entities are *gunmen, vessel, food*, and *ship*; and time-marking expressions include the aspect-marking *have*, the conjunction *after*, and the phrase *for more than two months*.

A dependency parse of the sentence is given as Figure 1; frames are associated with frame-bearing words; frame elements are associated with the dependent phrases in the diagram. The events identified in the narrative, in the order of their occurrence, are the tsunami, the chartering of a ship, the transporting of food aid (interrupted), the hijacking, the illegal retention of the ship, the ship's release, and the announcement about the release. The meaning of *after* shows that the releasing followed the holding; the fact that the reported event is *have released* rather than *will release* shows that the releasing event preceded the reporting event; since the vessel was carrying food aid at the time of the hijacking, the chartering and launching of the rescue vessel preceded the hijacking. Since in principle one can carry food aid for potential victims of a predicted tsunami, and it is not a **linguistic** fact that one does not charter a ship that is already on a mission, the ordering of these sub-events cannot be determined on linguistic grounds alone.

The category of event-introducing frame-evoking words is not limited to verbs. To be a victim is to be a participant in some unfortunate event, and this event is generally expressed as a modifier of the noun: *an X victim*, or *a victim of X*. The noun *tsunami* names an event which itself has no obligatory syntactic dependents, though it can have qualifiers indicating location, time, intensity, etc. In addition to expressions that directly point to the event that produced victims (the X in the above phrasings), it is possible to interpret the word *victim* alone as implying the existence of such, in a context in which information about the causing event is recoverable nearby. Thus, whenever we encounter a sentence like *Were there any victims?* it must be the case that mention of the mishap can be found in a recent part of the discourse.

There are many linguistic ways in which information about eventualities and claims about their actuality can be presented in a sentence. These include, most straightforwardly, frame-evoking verbs and nouns, and the ways in which the phrases they are in **grammatical construction** with contribute to "populating" ("filling the slots of") the frames. Verbs whose complements (including direct objects) are understood as participants in the frames they evoke are amply illustrated in the Hijacking text: the UN says the proposition about the ship's release; the gunmen hijacked the ship; the gunmen released the ship; the ship carried food aid; the hijackers held the ship for several weeks.

Information about reliability and source-attributions of event reports can be provided in meta-data and in text-internal mention of evidential sources. The Hijacking text itself is attributed (in the call for papers) to the Voice of America: knowing the mission of the VOA might be relevant to evaluating the reliability of the report, but that is not a linguistic matter. The reporting of the hijacking and releasing incident is attributed to the United Nations: again, evaluating the truth of the report by considering its source is not a linguistic matter.

FrameNet Annotation

(2) the United Nations SAYS **Statement** Somali GUNMEN **Bearing_arms** who HIJACKED **Piracy** a U.N. - CHARTERED **Renting** VESSEL **Vehicle** CARRYING **Bringing** FOOD Food AID **Assistance** for tsunami VICTIMS **Undergoing** have RELEASED **Releasing** the SHIP **Vehicle** AFTER **Time_vector** HOLDING **Inhibit_movement** it for more than two **MONTHS** Calendric_unit .

Space limitations will not allow a complete listing and discussion of the FN annotation of the Hijack sentence, but Ex. (2) shows all the currently-annotated frame-bearing words (in uppercase) followed by the names of their frames (in boldface). This includes a number of nouns such as *vessel* and *ship* in the **Vehicle** frame whose frames are relatively simple, and so will not be discussed further.

(3)
a. [**Speaker** The United Nations] SAYS [**Message** Somali gunmen who hijacked a U.N. - chartered vessel carrying food aid for tsunami victims have released the ship after holding it for more than two months] .
b. The United Nations says [**Agent** Somali gunmen who hijacked a U.N. - chartered vessel carrying food aid for tsunami victims] have RELEASED [**Theme** the ship] [**Time** after holding it for more than two months] .
c. The United Nations says [**Agent** Somali gunmen who hijacked a U.N. - chartered vessel carrying food aid for tsunami victims] have released the ship after HOLDING [**Theme** it] [*Duration* for more than two months] .
d. The United Nations says [**Perpetrator** Somali gunmen] [**Perpetrator** who] HIJACKED [**Vehicle** a U.N. - chartered vessel carrying food aid for tsunami victims] have released the ship after holding it for more than two months .

Ex. (3) shows the text with the names of FEs (in boldface) for some of the more interesting frames: (3) a. shows the FEs for the top-level predicator, *says*, which comprise the entire sentence. (3) b. shows the FEs for the next highest, *release*; note that, *release* together with its FEs covers the CONTENT FE of the higher verb *says*, meaning that the main CONTENT of the U.N. statement has to do with releasing (and is in a frame of the same name). Some of the FE names

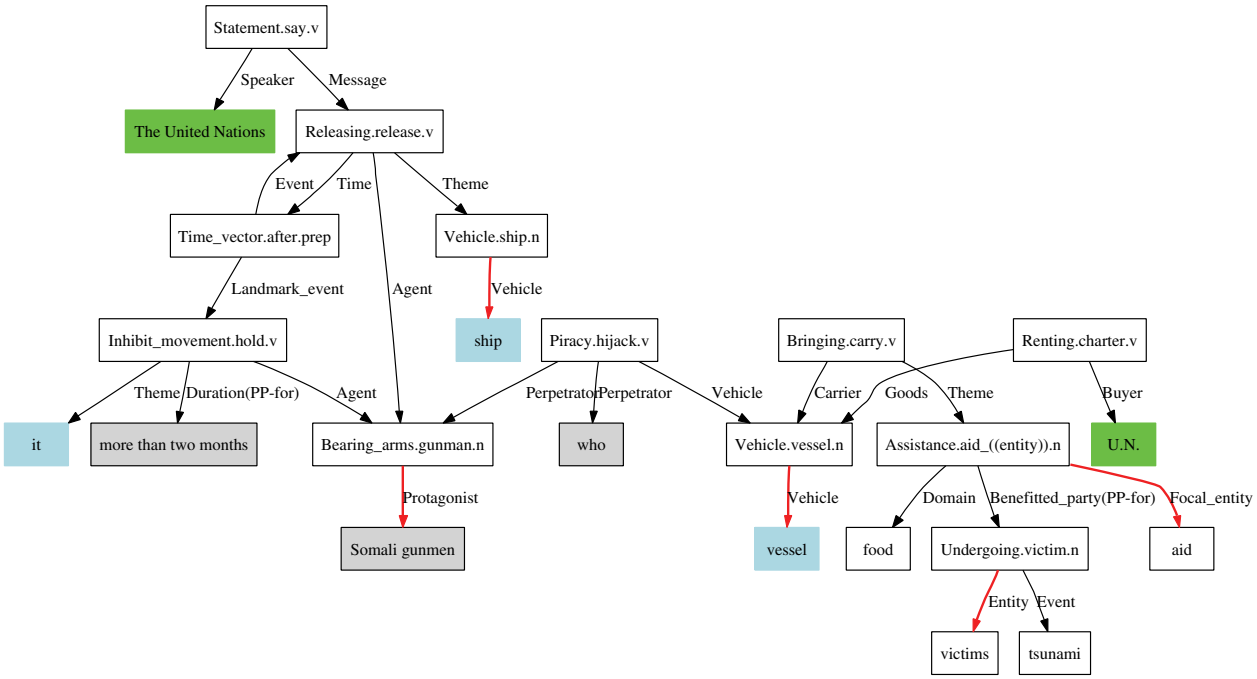

Figure 1: Frame Dependency Graph of Hijack Sentence

are quite specific as in Ex. (3) d .; e.g., since *hijack* evokes the **Piracy** frame, the thing hijacked is not just a THEME (as in Ex. (3) b. and c.) but a VEHICLE.

FrameNet-annotated texts like these can be produced both as XML (which is more human-readable but more project-specific) and as RDF-OWL, which is almost impossible for humans to read, but tractable with general-purpose inferencing tools. A more human-readable representation may be the dependency graph shown in Fig. 1, produced by means of the graphics package GraphViz.

In this figure, frame-bearing words are represented as nodes with their FEs as their dependents; the text of the node itself is <Frame name>,<LU name>.<POS>. The arrows to the dependents are labeled with the FE name (and also in two instances, with the phrase type PP *for*). The anaphora chains are indicated by colors on the nodes, blue for *it–ship–vessel*, and green for *United Nations* and *U.N.*. The red arrows pointing to the lower nodes labeled *ship, vessel, victims*, and *aid* represent cases in which the frame-evoking noun also denotes the filler of one of the FEs; this is common with non-event nouns. All of the information shown in this graph was extracted algorithmically from the XML format of the FrameNet annotations.

How much of the meaning of the sentence is captured in the dependency graph? Simple inspection of the graph shows that there are two clusters of nodes, one connecting an Aiding event and a vessel and the other connecting gunmen to events of Holding and Releasing. For each of the events in each cluster, we have identified the participants: the U.N. as the Lessee of the Renting event, the tsunami victims as the Benefitted_party of the Assistance_aid event, etc.[1] The event which connects the two clusters is an instance of Piracy evoked by the word *hijack*. To express more detail about the temporal and causal connections between these events, we must use a more powerful means of representation, such as X-Schemas, discussed in the next section.

From frames to inference

FrameNet frames provide the basis for identifying the roles of various events and a means to identify the various participants of the events (role fillers) from the text. However, there remains a gap between the FrameNet analysis of the input text and the kinds of event related inferences that can be made from the text. In previous work (Narayanan & McIlraith 2002; Chang, Narayanan, & Petruck 2002; Sinha & Narayanan 2005), we have addressed this state of affairs and have come up with a way to link FrameNet frames with a rich model of event structure that simulates the various events and their interactions. While details of our mapping and of the event simulation model are outside the scope of this paper, the central idea is that FrameNet frames index into parameterized domain models that capture important relationships between events.

FrameNet frames and frame element bindings provide parameters for event simulations which capture structural relationships among participants in a dynamic scenario. Sim-

[1] Note that the Statement event, which is highest in the syntactic parse tree, is in fact peripheral to everything else.

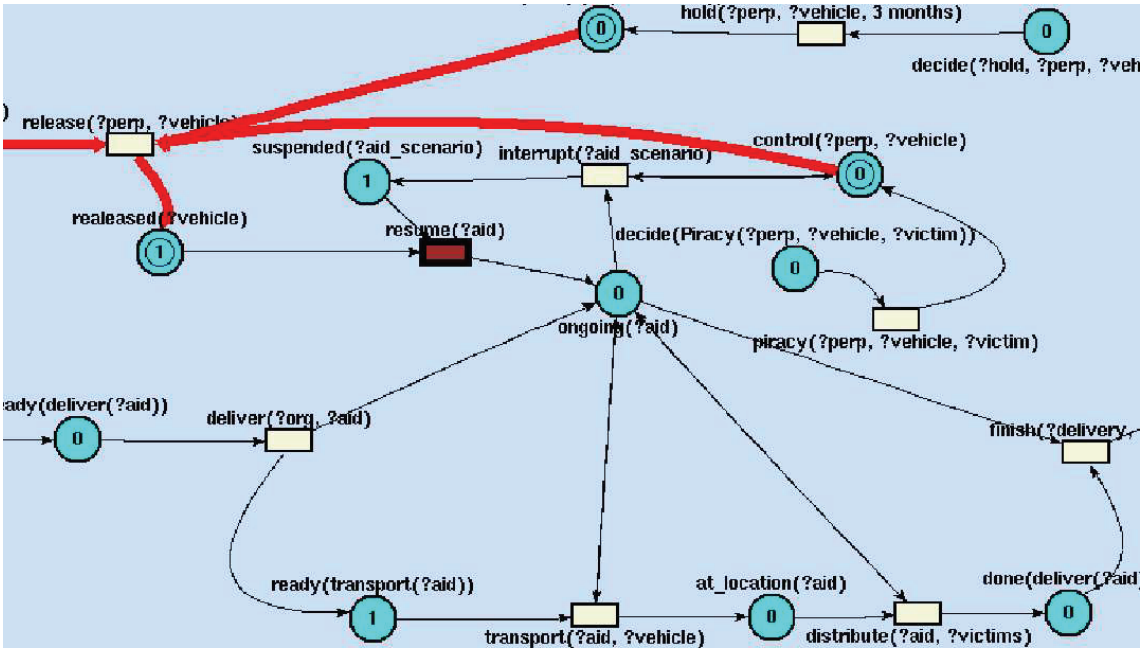

Figure 2: Inference with an Event Model. Shown is a fragment of the model of the **Assistance_aid** interacting with the model of the **Piracy** frame. Frames extracted from the input provide an index into a parameterized event model and supply bindings and parameters to an event simulation. Shown here is the simulation of the events. The event in question is the release stage of the **Piracy** frame (as specified by the Perfect aspect in the input narrative).

ulations of the events and their interactions model the evolution of the scenario and inference procedures are able to go backwards and forwards in time for predictions, possible event trajectories, and participant bindings. This formalism provides a flexible means of accounting for linguistic perspective and other inferential effects of processing the input text.

Figure 2 shows the event model evoked by the various frames in the input text. From the dependency graph, there are the two main interacting scenarios. The first is that of an organization providing assistance or aid to victims (of the tsunami). The model of this scenario involves several coordinated events and sub-events such as the obtaining of a vehicle and the aid, transporting the aid to the location of the victims, delivering the aid to the appropriate locations etc. This model is indexed from the **Assistance_aid** frame in input, and a second event corresponds to the **Piracy** which involves gaining control of a vehicle and holding it for a period of time and possibly releasing it after. Figure 2 models the interactions of these scenarios providing the basis for important inferences from the text:

1. The Piracy event is in fact an *interruption* to the ongoing aid delivery. The effect of the event is that the vehicle is under the control of the perpetrator and that the aid delivery remains *suspended*. This information comes from the FrameNet frame element bindings for the roles of the Piracy frame (evoked by the word *hijack*).
2. Being in control of the vehicle enables the perpetrator to hold the vehicle for a period of time (in this case 2 months). Meanwhile the aid delivery is *suspended*, since the vehicle remains in control of the perpetrator.
3. When the perpetrator has control of the vehicle (a precondition), he can release the vehicle, which lets the aid delivery *resume* and so the state goes back from being *suspended* to being ongoing.
4. The stage being referred to in the narrative is the completion (from the perfect aspect of *have released*) of the release stage of the hold and release scenario, which is a precondition to the resumption of aid. This is the stage shown in Figure 2.

It is this fine-grained interaction that enables various important event related inferences to felictously occur, such as a) the aid was ongoing and the food was being transported to the destination while the hijacking occurred, b) the hijacking suspended the delivery of aid and the victims could not get the aid for the time that the perpetrators held it, c) the perpetrators have released the vessel; the UN assistance and aid delivery can resume at this point, d) at some time in the future the food may reach its destination and (if the distribution event can take place) the victims of the tsunami can receive aid, etc. All these inferences are formal procedures (reachability, MAP estimation) on the graphical model of the events.

Beyond the Hijacking Text

Beyond *says*: Verbs with sentential complements

The Hijacking text has only one source-attributing expression, *The United Nations says*, but there are numerous ways of introducing a proposition that vary in respect to the speaker's commitment to the truth of the proposition. A simple statement like *Saddam Hussein recently sought significant quantities of uranium from Africa* expresses the speaker's commitment to the truth of the proposition. A report of such a proposition of the form *British intelligence has **reported** that Saddam Hussein recently sought significant quantities of uranium from Africa* does not imply such a commitment. President Bush's famous statement *British intelligence has **learned** that Saddam Hussein recently sought significant quantities of uranium from Africa* **does** imply that the president was certain of the truth of the claim. Later White House paraphrases of his statement, in which *learned* was replaced with *reported*, were introduced to withdraw that commitment. The issue is one of recognizing the indirect assertorial force of language that is not itself dedicated to making assertions.
[2]

Beyond Verb+Object I: Verbs with verb-headed complements

Except for the single verb that takes a clausal complement (*says*), all of the major verbs in the Hijacking text are simple transitives: the UN *charters* a vessel, the vessel *carries* food aid, the gunmen *hijack* the vessel, the gunmen *hold* the vessel, the gunmen *release* the vessel. In this section and the next we will consider constructions of which there are no examples in the Hijacking text, such as verb-to-verb structures (*refuse to leave*) and verb-to-event-noun structures (*make a promise*).

Verb-to-verb constructions of the "control" kind, typically associated with infinitive phrase complements, differ in respect to questions of evidentiality, factivity and event-counting. The sentence *The bodyguard **tried** to **poison** the emperor* introduces two events, an actual Trying event and a defeasible Poisoning event; but a structurally similar sentence like *The bodyguard **managed** to **poison** the emperor* introduces only one event, Poisoning, and affirms its actuality. In *He **claimed** to have poisoned the emperor* we have two events, the Claiming event being asserted, the Poisoning event not-asserted; in *He refused to poison the emperor*, again, we have two events, with the non-actuality of the second one highly implicated. In other words, structures of roughly the same syntactic form can differ in a number of ways in these regards, and such information can be coded among the combinatorial affordances of the "higher" verbs, i.e., *try, manage, pretend, claim, refuse*, and many scores of others.

FrameNet has many frames dealing with control verbs and related nouns, adjectives, etc., such as Success_or_Failure (*fail, manage, pull off, succeed, success, successful, unsuccessful*); note that these include both factives and conterfactives, as shown by the name of the frame. At present, the necessary LU-by-LU information is not normally coded in the database. However, there is a precedent for how this can be handled; in several frames related to judgment, LUs representing positive judgment (*respect, praise*) and negative judgment (*despise, criticize*) are treated in the same frame, but distinguished with regard to "positive" or "negative" by the use of a (FN-specific) semantic type on the LUs. A similar approach should work in the case of factivity.

Beyond Verb+Object II: Verbs with complements headed by frame-bearing nouns

We now turn to verbs with noun-headed complements in which the noun designates an event. The questions to ask for a verb-noun pairing of this kind are Which is the semantic head (the verb or the noun)?, How many events are encoded?, and Do the verb and the noun share arguments?.

In certain cases we clearly have two events and there is no necessary sharing of their participants. In a sentence like *Joan **overheard** my broker's **advice** to me* we have the Advising event and the Overhearing event, and they have no shared participants. In *My broker **gave** me good **advice***, there is one event, and the verb give serves merely to enable the frame to be expressed verbally. The first two arguments of *gave*, *my broker* and *me*, are the participants in the Advising event.

There is a very large class of structures of this kind, the most common of them using verbs like *make, have, give*, and *take* (*make progress, have a fight, give advice, take a bath*), but also many that are more case-specific, including examples in which the event-noun is not a derivative of a verb: *say a prayer, wage war, wreak havoc*. In all such cases there is one event, and the syntactic arguments of the verb are shared with the arguments of the noun.

We can refer to these as **support constructions**; they include not only those of the pattern V+N just illustrated, but also V+N+Preposition (*take advantage of, take possession of, pay attention to*) and V+Preposition+N (*put into effect, take into account, take under consideration*), and a few other patterns.

The choice of support verb can also express (1) a "perspective" on an event, by highlighting one participant's role (*perform an operation* and *undergo an operation*, *inflict injury* and *sustain an injury*, *give advice* and *get advice*) or (2) differences in register (*make a complaint, lodge a complaint, register a complaint*, and *file a complaint*). Here too the verb-noun combination still stands for only one event.

Beyond these simple cases are situations in which the verb represents some desired or undesired outcome of the event designated by the noun. To *get advice* is to be a participant in a single event, but to *take* (someone's) *advice* is to participate in two events, first the Advising event and then the

[2] The issue of factivity arises mainly in the domains of cognition and communication, where FrameNet is fairly well-developed, but at present, very little information as to factivity is available in the FrameNet database. There are some frames, such as Opinion, in which none of the LUs are necessarily factive. In general, however, factivity or non-factivity of content can vary within a frame and would have to be marked LU by LU; e.g. in the Coming_to_believe frame, *find out, learn* and *realize* imply that the Content is true, while *conclude, deduce, guess* and *infer* carry no such implication.

Uptake. To *make a promise* is to participate in one event; to *keep a promise* (or to *break a promise*) is to participate in two events, first the initial Promising event, and then the event of fulfilling (or failing to fulfill) the conditions of the promise.

Beyond Overt Anaphora: Definite and Indefinite Zeroes

The Hijacking text has two anaphoric chains, the first involving two graphic forms of a single name (United Nations, U.N.) the second involving an indefinite NP (a vessel), a co-referring definite NP using a synonym (the ship), and a pronoun (it). In particular, there are no instances in which entities or events are introduced indirectly, as "missing complements" of specific lexical items. A superficial look at a sentence like *Witnesses reported that the suspect was wearing a beret*, would suggest that there are only two events, Reporting and Wearing. The speaker role in the Reporting event is held by the witnesses, and the participants in the Wearing event are the suspect and the beret. We know, however, that there had to be a particular event–mentioned elsewhere in the text—**to which** the witnesses were witness, and there had to be a criminal act, mentioned elsewhere in the text, of which the suspect was a suspect. It is important to realize that this knowledge is not "simply" reasoning based on world knowledge: it is not that we know about witnesses that they have observed something salient, but that the word *witness* **denotes** someone who has just this role; it is not that we know about suspects that people think they might be responsible for some criminal act, but that the word *suspect* **denotes** someone in just this position. A context-free, conceptually complete use of *witness* would indicate the related incident: *witness to the accident*; a context-free conceptually complete use of *suspect* would indicate the related criminal act: *suspect in the incident*.

A more complex example of the same phenomenon can be seen in the following sentence: *When other people arrived at the scene, and saw that there was no-one in the burnt-out cockpit, they assumed that the pilot had bailed out.* How is it that we are quick to conclude that there has been a plane crash and that there were earlier visitors to the scene where this took place. In this sentence the words with unmentioned semantic complements include *other, scene, cockpit, pilot* and *bail out*. In the case of *other*, there is a missing *than*-phrase, standing for 'other than the person or people just mentioned'; the word *scene*, perhaps especially in the phrase *at the scene*, typically entails an understood *of*-phrase indicating some event, most frequently a mishap; the noun *pilot* denotes someone who flies a plane (again, it is not because we know that pilots are the kind of people who fly planes, but that the word *pilot* is dedicated to naming occupants of such a role); and one meaning of the verb *bail out* is 'to jump out of an airplane'.

We believe that this sentence demonstrates that the kind of anaphora resolution that is typically carried out as determining a relation between "mentions" of the same entity needs to be expanded in order to take on a new kind of representation: one which indicates, for particular frame-evoking words or structures, the kind of thing that has to be recoverable in the nearby context. There is a great deal in understanding this sentence - if given no other context! - that does involve world knowledge. But most or all of the information that abductive reasoning could construct from this sentence alone can be found in the previous part of the text itself. The abridged version of the text from which the example was extracted is:

> One July day in 1943 at Blaconsthorpe, Norfolk, a born-deaf farm labourer named Kenneth Andrews . . . saw a R.A.F. Typhoon fighter plane crash in flames. . . . Andrews ran to the downed plane and saw that the pilot was trapped in his cockpit.
>
> Without any regard for his own safety, Andrews climbed onto a wing of the blazing plane, and kicked frantically at the toughened perspex of the cockpit canopy until it shattered, then dragged out the semi-conscious pilot. Seconds after Andrews had dragged away the pilot, the plane exploded, and was reduced to a smouldering wreck.
>
> Andrews flagged down a passing vehicle, and saw that the driver would take the pilot to the nearest hospital, then calmly went home with his wife to have his dinner. **When other people arrived at the scene, and saw that there was no-one in the burnt-out cockpit, they assumed that the pilot had bailed out.** . . .

Thus FrameNet can provide sentences annotated with tags indicating what sorts of fillers are to be found in the surrounding contex. If anaphora resolution systems can exploit these, a new level of understanding of text coherence may be attainable.

Acknowledgments

This work was supported in part by the Advanced Research and Development Activity (ARDA)'s Advanced Question Answering for Intelligence (AQUAINT) program.

References

Chang, N.; Narayanan, S.; and Petruck, M. R. 2002. Putting frames in perspective. In *Proceedings of 19th International Conference on Computational Linguistics*. Taipei: COLING.

Fontenelle, T., ed. 2003. *International Journal of Lexicography*, volume 28. Oxford University Press. (Special issue devoted to FrameNet.).

Narayanan, S., and McIlraith, S. 2002. Simulation, verification and automatic composition of web services. In *Proceedings of the Eleventh International World Wide Web Conference (WWW2002), Hololulu, Hawaii, May 7-10, 2002*.

Sinha, S., and Narayanan, S. 2005. Model based answer selection. In *Proceedings of the Workshop on Textual Inference, 18th National Conference on Artificial Intelligence*. PA, Pittsburgh: AAAI.