Precedes: A Semantic Relation in FrameNet

Miriam R. L. Petruck and Gerard de Melo
International Computer Science Institute
Berkeley, California, USA
miriamp@icsi.berkeley.edu, demelo@icsi.berkeley.edu

Abstract

Automatic language processing systems depend on, among others factors, the effectiveness in modeling human cognitive abilities, including the capacity to draw inferences about prototypical or expected sequences of events and their temporal order. Appropriate response to a crisis is as important for public security as are efforts to prevent any such natural or man made disaster. Recent research (Mehrota et al. 2008) has recognized the need for accurate and actionable situation awareness during emergencies, where timely status updates are critical for effective crisis management. The present paper constitutes a contribution to situation awareness for Natural Language Processing (NLP) applications to improve communication among first responders, and features the frame-to-frame semantic relation **Precedes**, as implemented in FrameNet (http://framenet.icsi.berkeley.edu). Specifically, this work demonstrates the necessity and importance of the information encoded with **Precedes** for NLP applications, advocating the inclusion of such information in systems for security applications.

Keywords: FrameNet, semantic relations, inferencing

1. Introduction

The success of any automatic language processing system depends on, among others factors, its effectiveness in modeling human cognitive abilities, including the capacity to draw inferences about prototypical or expected sequences of events and their temporal order. Appropriate response to a crisis is as important for public security as are efforts to prevent any such natural or man made disaster. Recent research (Mehrota et al. 2008) has recognized the need for accurate and actionable situation awareness during emergencies, where timely status updates are critical for effective crisis management. The present paper constitutes a contribution to situation awareness for Natural Language Processing (NLP) applications to improve communication among first responders, and features the frame-to-frame semantic relation **Precedes**, as implemented in the FrameNet database (http://framenet.icsi.berkeley.edu).

Based on the principles of Frame Semantics (e.g. Fillmore, 1975; Fillmore, 1977; Fillmore, 1985) and committed to documenting its findings through corpus attestations,[1] FrameNet is an ongoing lexical resource development project that documents the **valences**, or semantic and syntactic combinatorial possibilities, of English vocabulary. With computer-assisted annotation of examples and automatic tabulation and display of the results, FrameNet records the valences for each word in its lexicon. The project's main product is its lexical database, currently containing over 1,100 frames, nearly 12,500 lexical units, and approximately 190,000 examples. Befitting of the larger endeavor, the database includes frame-to-frame relations for its hierarchy of semantic frames.

Much of the written work and public presentations of the FrameNet team and its affiliates have attended to the most important frame-to-frame relations recorded in the database, **Inheritance**, **Subframe**, and **Using** (Fillmore and Baker, 2004; Fillmore et al., 2004). Petruck et al. (2004) characterized the lexicographic imperative for adding **Inchoative_of** and **Causative_of** to the inventory of semantic relations that FrameNet records. Chang et al. (2002) provided a structured event formalism that translates FrameNet's informal descriptions into a representation appropriate for simulative inference, also offering a means of handling linguistic focus, which FrameNet captures with **Perspective_on**.[2] The present paper focuses on **Precedes**, the semantic relation in FrameNet that captures the notion of temporal ordering, and demonstrates its necessity for natural language understanding (e.g. Buchardt et al., 2009; Shen and Lapata, 2007; Fillmore and Baker, 2001).

2. Frame Semantics

At the heart of Frame Semantics is the **semantic frame**, a schematic representation of an event, object, situation or state of affairs whose **frame elements** (**FE**s) identify participants and props and whose underlying conceptual structure speakers access for both encoding and decoding purposes. The semantic frame, parts of which are indexed by words that *evoke* the frame, is a cognitive structuring device used in the service of understanding (Fillmore, 1985). FrameNet distinguishes three categories of FEs: **core**, **peripheral**, and **extra-thematic**. Core FEs are frame specific and uniquely define a frame, capturing conceptually necessary aspects of the scene. Peripheral FEs identify characteristics of situations and events more generally, including the time or place of

[1] FrameNet primarily uses the British National Corpus, and to some extent the American National Corpus.

[2] Chang et al. (2002) foreshadowed the introduction of **Perspective_on** into FN. **Perspective_on**, a refinement of **Using**, first appeared in data release 1.3 (2006).

an event, as well as the manner in which an event occurs. Extra-thematic FEs situate an event or state of affairs against the backdrop of another event or state of affairs, such as the frequency with which an event occurs, or a description of a participant in an event in terms unrelated to the event. Conceptually, extra-thematic FEs are not part of the frame in which that type of FE appears, instead belonging to more abstract frames where they fill argument roles of their own.[3] A Frame Semantics description of a word identifies the frame or frames that constitute the conceptual basis of a given sense, and specifies the ways that structures headed by the word realize those FEs.

FrameNet defines `Avoiding` as a situation in which an AGENT avoids an UNDESIRABLE_SITUATION under certain CIRCUMSTANCES, where that situation may be an event or an activity.[4] Whereas AGENT and UNDESIRABLE_SITUATION are core Frame Elements, CIRCUMSTANCES is a peripheral FE, since it characterizes an aspect of a wide range of events in addition to avoiding. Following Cruse (1986), FrameNet adopts the **lexical unit** (**LU**) as the focus for lexicographic annotation, defining an LU as a pairing of a lemma and a frame. Among the LUs that figure in `Avoiding` are *avoid*, *avoidance*, *evade*, and *evasion*.

Example (1) illustrates the LU *avoid*.v, which evokes the `Avoiding` frame, annotated with respect to that target LU, along with the triples of information that FrameNet records for each FE, including phrase type (PT) and grammatical function (GF).

1. [AGENT The reporter NP/External] **AVOIDED** [UNDESIRABLE_SITUATION entering the roped off area VPing/Dep].

The NP *the reporter*, instantiating the AGENT has the GF External; and the VP*ing entering the roped off area* realizes the UNDESIRABLE_SITUATION, functioning grammatically as a Dependent (Dep).[5] A FrameNet lexical entry provides a table of the valence patterns, or combinatorial possibilities, specifying the mapping of semantic roles to syntactic structures and showing the full array of syntagmatic relations for that word. Below, Figure 1 shows a partial valence table for *avoid*.v in `Avoiding`, displaying only the core FEs AGENT and UNDESIRABLE_SITUATION.

Example (2) illustrates FrameNet annotation for a sentence that also realizes non-core FEs, here TIME and CIRCUMSTANCES, the former as PP/Dep and the latter as Sinterrogative/Dep (PT/GF).

2. [TIME At the beginning PP/Dep] [AGENT the reporter NP/External] **AVOIDED** [Undesirable_situation entering the roped off area VPing/Dep] [Circumstances while looking for evidence Sinterrogative/Dep].

FrameNet distributes the valence tables for each lexical entry in XML, making this syntagmatic information accessible for use in NLP applications.

Number Annotated	Patterns	
(140) TOTAL	Agent	Undesirable_situation
(15)	CNI --	NP Ext
(17)	CNI --	NP Obj
(2)	CNI --	VPing Dep
(89)	NP Ext	NP Obj
(1)	NP Ext	PP[including] Dep
(3)	NP Ext	Sing Dep
(9)	NP Ext	VPing Dep
(1)	NP Obj	VPing Dep
(3)	PP[by] Dep	NP Ext

Figure 1: Partial Valence Table for `Avoiding`.*avoid*.v

3. Frame-to-Frame Relations in FrameNet

FrameNet records frame-to-frame relations in its hierarchy of semantically organized frames, also making that information available for natural language processing applications. Figure 2 depicts the relevant frame-to-frame relations that FrameNet has recorded for the `Employment_scenario` frame.

Inheritance exists between a parent frame and a child frame under specific circumstances: for each FE, frame relation, and semantic characteristic in the parent, the same or a more specific corresponding entity in the child exists, as in the relationship between `Employment_end` and `Firing`. **Using** is a relationship between a child frame and parent frame in which only some of the FEs in the parent have a corresponding entity in the child; if such exist, they are more specific. Using holds between `Fields` and `Employment_scenario`, where the FEs ACTIVITY, PRACTITIONER and WORK in `Fields` are the more specific instances of TASK, EMPLOYEE and POSITION in `Employment_scenario`, respectively. FrameNet uses **Perspective_on** (Chang et al. 2002) to distinguish between neutral and *perspectivized* frames, the latter identifying different points of view of other participants in the larger scenario. As a consequence, whereas `Employment_scenario` is neutral in terms of participant point of view, `Employer's_scenario` and `Employee's_scenario` capture the perspective of the employer and employee, respectively. The **Subframe** relationship characterizes the different parts

[3] See Chapter 3, Ruppenhofer et al. (2010), for a detailed discussion of FEs and coreness.

[4] The paper uses the following typographical conventions: **bold-face** for special terms in the prose; small caps for FEs; bold-face caps for targets; and Courier New for frame names.

[5] FrameNet uses a limited set of grammatical functions. See Chapter 5, Ruppenhofer et al. (2010) for more information.

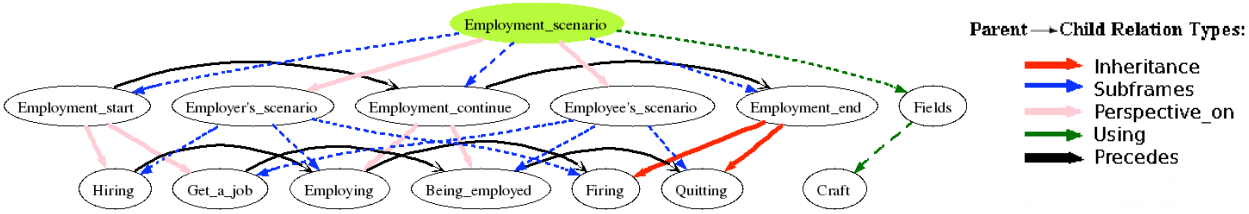

Figure 2: The Employment_scenario and its Frame-to-frame Relations

of a complex event in terms of the sequences of states of affairs and transitions between them, each also describable as a frame. **Precedes** captures the temporal ordering of subevents within a complex event. The relation holds between component subframes of a single complex frame, and provides additional information to the set of **Subframe** relations, as in `Hiring`, `Employing`, and `Firing`, each a separate frame in the complex `Employer's_scenario`. **Precedes** is also the only relation that allows cycles, as for example with repeated hirings, employings, and firings.[6]

4. Fire Fighting: Search and Rescue

Fire fighters are among the first to appear on the scene of an emergency, as occurred on September 11, 2001. Here, we generally characterize the complex Fire Fighting (FF) scenario and then focus on Search and Rescue, a sub-phase of FF, to demonstrate the necessity and importance of **Precedes** for NLP applications.

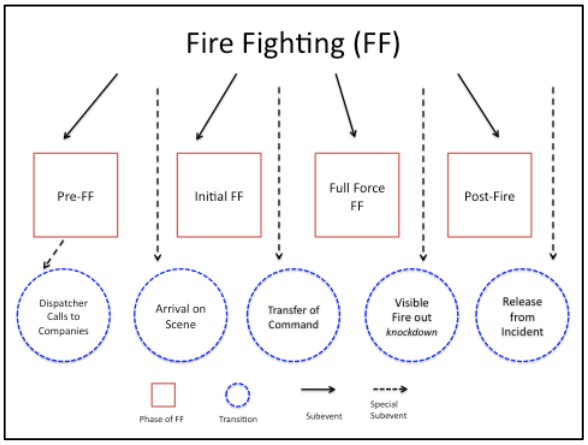

Figure 3: Fire Fighting Phases and Transitions

Figure 3 depicts the major phases and transitions in FF, where FrameNet would define each phase as a semantic frame. Although not surprising given the situation, note that transitions between the major phases of FF are primarily communication events, all of which follow Chain-of-Command conventions.

Figure 4 captures the sub-phases of Full Force Fire Fighting (F^4), in FrameNet terms subframes of F^4. As Figure 4 suggests, during Full Force Fire Fighting, numerous actions or sub-events occur simultaneously, one of which is Search and Rescue. But for Salvage, an ongoing activity during Full Force Fire Fighting, Search and Rescue takes precedence over other activities, fire conditions permitting, and continues until the unit determines that no one else needs to (or can) be rescued.

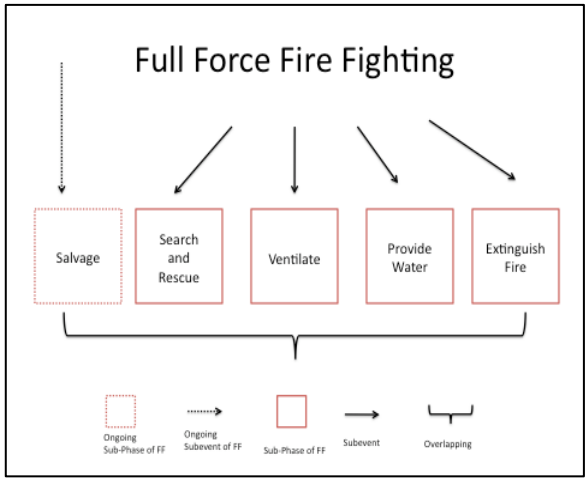

Figure 4: Full Force Fire Fighting

Fire fighters must communicate their status to the incident commander (IC), usually a battalion chief, so that the IC can manage the situation effectively. For instance, members of the unit that conduct Search and Rescue, typically medics, report their whereabouts and activity to the IC. Figure 5 depicts the subparts that constitute the more complex Search and Rescue scenario, each a separate subframe, and displays the relation Precedes holding between the relevant subframes.

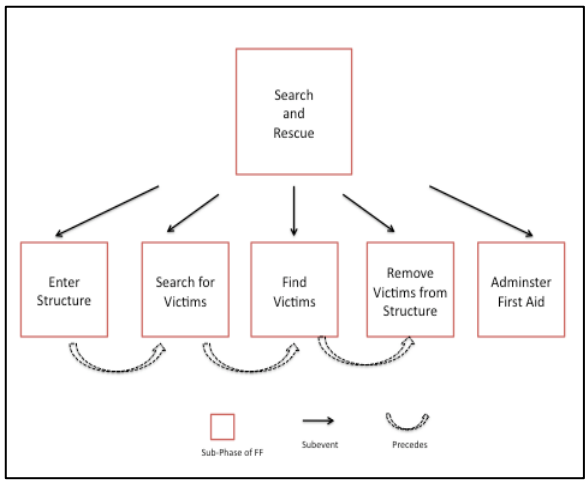

Figure 5: Search and Rescue

[6] Petruck et al. (2004) describes how FrameNet records the relations **Causative_of** and **Inchoative_of** to separate causative events, inchoative events, and statives, with each type of predicate in its own frame.

Consider (3), a possible utterance of a fire fighter conducting Search and Rescue, also updating the IC. Identifying the structured knowledge that English speakers must access to understand the entire utterance will highlight the need for **Precedes** in NLP applications. The clipped nature of (3) is more typical of speech during crisis response than of written text, thus (3) serves as a reminder that participants in Search and Rescue must exploit a wealth of shared background knowledge for effective communication.

3. Helped woman who was gasping.
   Inside the stairwell.
   Lots of smoke.
   Heading out now.

Helped evokes a scene in which someone administers first aid to a victim; and the noun phrase *woman who was gasping* describes that victim. The adverb *inside* implies that the participants are in an enclosed structure; and *the stairwell* evokes that structure. The only word that provides any hint of Fire Fighting is *smoke*. While humans can infer that the speaker of (3) entered a building, searched for, and found a gasping woman *before* administering aid, NLP systems cannot unless the system encodes information about the expected order of events in Search and Rescue. **Precedes** captures the chronological order of events in the subframes of Search and Rescue, independent of their place in the utterance's word order. The frame-to-frame information in Figure 5 allows the automatic decision about the order of events in (3). Absent **Precedes**, and other knowledge structured in the FF scenario (some of which we described), the word order of LUs in (3) could lead a system to conclude incorrectly that the main events unfolded as follows: Administer Aid, Gasp, Enter, and Exit. Because saving lives is the highest priority in Search and Rescue, Figure 5 does not show **Precedes** between Remove Victims from Structure and Administer First Aid since only the fire fighter conducting Search and Rescue can determine which order is best.

5. Conclusion

Frame Semantics is among the most useful techniques for deep semantic analysis of linguistic material, primarily text. This paper has illustrated the contribution of Frame Semantics, as instantiated in FrameNet, to research on situation awareness, highlighting the role of the relation **Precedes** for natural language understanding during crisis response. Except for Ruppenhofer et al. (2010), **Precedes** has received virtually no attention in FrameNet-related research. This work fills that gap by demonstrating the necessity of the information encoded with **Precedes** for NLP applications, advocating the inclusion of such information in systems for security applications.

Currently, FrameNet includes 82 instances of Precedes, only 4.9% of the frame-to-frame relations recorded.[7] As FrameNet continues to expand and cover more areas of English vocabulary, necessarily defining frames that characterize a greater number of complex event and state scenarios, instances of the **Precedes** relationship will increase also. Moreover, to enhance its usefulness, FrameNet must add other relations to its repertoire of frame-to-frame relations. For instance, Hasegawa et al. (2011) proposed introducing two relations new to FrameNet, i.e. **Symmetric_antonymy** (*male*/*female*) and **Asymmetric_antonymy** (*love*/*hate*) to capture different types of negation that may hold between certain LUs, enriching the FrameNet database, and facilitating its use as a resource for paraphrasing. Somewhat comparably, to facilitate inferencing work, others have suggested that FrameNet implement an entailment relationship in the database (Ovchinnikova et al. 2010).

[7] The Appendix lists examples in FrameNet, where **Precedes** holds between a frame in column 1 and its partner in column 2.

6. Acknowledgements

DARPA Contract award #HR0011-11-C-0017 to the International Computer Science Institute in Berkeley, CA, supported part of the work presented in this paper. The authors are indebted to Charles J. Fillmore, Jim Hieronymous, Collin F. Baker, Michael Ellsworth, and Russell Lee-Goldman for helpful comments. Only the authors are responsible for the final version.

7. References

Berlin, B., D.E. Breedlove and P.E. Raven (1973). General Principles of Classification and Nomenclature in Folk Biology, *American Anthropology*, 75, pp. 214–42.

Brown, C.E., J. Kolar and B.J. Torrey, T. Truoong-Quang and Volkman, P. (1976). Some general principles of biological and non-biological folk classification. *American Ethnologist* 3, pp 73–85.

Burchardt, A, M. Pennacchiotti, S. Thater and M. Pinkal. (2009). Assessing the impact of frame semantics on textual entailment. *Natural Language Engineering* 15(4), pp. 527–550.

Chang, N., S. Narayanan and M.R.L. Petruck (2002). Putting Frames in Perspective. In *Proceedings of the Nineteenth International Conference on Computational Linguistics*, Taipei: COLING, pp. 148-154.

Cruse, A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Fillmore, C.J. (1975). An alternative to checklist theories of meaning. In *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, Berkeley: BLS, pp. 123-131.

--- (1977). The need for a frame semantics in linguistics. In H. Karlgren (Ed.) *Statistical Methods in Linguistics*. Stockholm: Skriptor, pp. 5-29.

--- (1978). On the organization of semantic information in the lexicon. In *Papers from the Parasession on the Lexicon*, Chicago: CLS, pp. 148-173.

--- (1985). Frames and the Semantics of Understanding. *Quaderni di Semantica*, 6(2), pp. 222-254.

Fillmore, C.J. and Baker, C.F. (2001). Frame Semantics for text understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop*, Pittsburgh: NAACL, pp. 59–64.

--- (2004). The evolution of FrameNet annotation practices. In *Proceedings of Building Lexical Resources from Semantically Annotated Corpora Workshop*. Lisbon: LREC, pp.1-8.

Fillmore, C.J., C.F. Baker, and H. Sato (2004). FrameNet as "Net". In *Proceedings of the Fourth International Conference on Language Resources and Evaluation*. Paris: ELRA.Volume IV, pp. 1091-1094.

Fillmore, C.J., M.R.L. Petruck, J. Ruppenhofer, and A. Wright (2003). FrameNet in action: The case of attaching, *International Journal of Lexicography*, 16(3), pp. 297-332.

Hasegawa, Y., R. Lee-Goldman, A. Kong and A. Kimi, (2011). FrameNet as a resource for paraphrase research. *Constructions and Frames*, 3(1), pp. 104–127.

Liston, J.L. (1972). The semantic structure of body part terms in Serbo-Croatian: I, The part-whole hierarchy. *Anthropological Linguistics* 14(8), pp. 323-338

McClure, E.F. (1975). Ethno-anatomy: The structure of the domain. *Anthropological Linguistics* 17.2: 78-88.

Mehrota, S., T. Znati and C.W. Thompson (2008). Crisis Management. *IEEE Internet Computing* 12(1), pp. 14-17.

Ovchinnikova, E., N. Montazeri, T. Alexandrov, J.R. Hobbs, M.C. McCord, and R. Mulkar-Mehta (2011). Abductive Reasoning with a Large Knowledge Base for Discourse Processing. In *Proceedings the Ninth International Conference on Computational Semantics*, Oxford: ACL pp. 225-234.

Petruck, M.R.L., C.J. Fillmore, C.F. Baker, M. Ellsworth and J. Ruppenhoffer (2004). Reframing FrameNet data. In *Proceedings of The Eleventh European Association of Lexicigraphy International Congress*, Lorient: EURALEX, pp. 405-416.

Ruppenhofer, J., M. Ellsworth, M.R.L. Petruck, C.R. Johnson, and J. Scheffczyk (2010). *FrameNet II: Extended Theory and Practice* (Web Publication http://framenet.icsi.berkeley.edu).

Shen, D. and M. Lapata. (2007). Using semantic roles to improve question answering. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague: ACL, pp. 12–21.

A. Appendix: Precedes holds from 1 to 2

#1	#2
Attempt	Success_or_failure
Existence	Ceasing_to_be
Get_a_job	Being_employed
Being_employed	Quitting
Hiring	Employing
Employing	Firing
Committing_crime	Criminal_investigation
Arrest	Arraignment
Birth	Death
Birth	Dying
Waking_up	Being_awake
Being_awake	Fall_asleep
Being_awake	Getting_up
Coming_to_be	Existence
Aiming	Hit_or_miss
Visiting	Visitor_departure
Invading	Conquering
Invading	Repel
Dying	Death
Confronting_problem	Resolve_problem
Ceasing_to_be	Out_of_existence
Fall_asleep	Sleep
Assemble	Meet_with
Arraignment	Trial
Notification_of_charges	Entering_of_plea
Entering_of_plea	Bail_decision
Change_of_phase	Altered_phase
Jury_deliberation	Verdict
Criminal_investigation	Criminal_process
Court_examination	Jury_deliberation
Text	Labor_product
Trial	Sentencing
Being_born	Dying
Being_born	Death
Employment_continue	Employment_end
Employment_start	Employment_continue
Visit_host_arrival	Visit_host_stay
Visit_host_stay	Visit_host_departure
Visiting_scenario_arrival	Visiting_scenario_stay
Visiting_scenario_stay	Visiting_scenario_departing
Visitor_arrival	Visiting
Event	Change_of_state_endstate
Sleep	Waking_up