

# Explaining the Structure of FrameNet with Concept Lattices

Francisco J. Valverde-Albacete\*

Dpto. de Teoría de la Señal y de las Comunicaciones,  
Universidad Carlos III de Madrid  
Avda. de la Universidad, 30. Leganés 28911. Spain

and  
International Computer Science Institute,  
1947 Center St. Berkeley, CA, USA  
`fva@tsc.uc3m.es`

**Abstract.** This paper reports on ongoing work to use Formal Concept Analysis as an auxiliary tool in understanding and visualising the wealth of data produced by lexical-resource building as embodied in the construction of FrameNet, a database to capture the syntax and semantics of language use in Frame Linguistics. We present proof of the abundance of concept lattices both in the theory of frames and in its present day incarnation, the FrameNet resource, with contributions that range from data-visualisation to the fine-tuning of some lexico-theoretical concepts better understood in terms of Formal Concept Analysis.

## 1 Introduction

FrameNet[1, 2, 3] is a lexical resource and database being developed at the International Computer Science Institute, ICSI. The name “FrameNet” reflects the twin facts that the project exploits the theory of Frame Semantics and that it is concerned with the semantics of networks through which word meanings are connected with each other. In this paper, we will be concerned with the modelling of Frame Semantics in terms of concept lattices and we will use the data in FrameNet to build concrete examples of such lattices.

The modelling of linguistic phenomena figures among the first applications of Formal Concept Analysis. Early references first applied the technique to the modelling of linguistic paradigms [4, 5, 6]. Lately the spectrum of applications has widened [7] for instance to the formalization of the sign relation [8]. In the lexical enterprise proper, a formalization of WordNet in terms of concept lattices has already been attempted [9, 10]; non-standard techniques like neighbourhood concept lattices have even been used to add structure to such a widely-used

---

\* This work has been partially supported by the Local Government of Madrid, Spain. grant CAM-07T/0014/2003-1 and by a grant for “Estancias de Tecnólogos Españoles en el International Computer Science Institute” of the Spanish Ministry of Industry.

resource as Roget’s thesaurus [11, 12] and even machine learning techniques have been used to induce the concept lattice of some lexical relations [13, 14]. We believe, however, that ours is the first attempt to model Frame Semantics with the help of Formal Concept Analysis techniques.

The unit of description in FrameNet is not the word of traditional dictionaries, or the Synsets of WordNet, but rather *frames*, the “schematic representations of the conceptual structures and patterns of beliefs, practices, institutions, images, etc. that provide a foundation for the meaningful interaction of a given speech community” [3]. The central idea of the whole frame semantic enterprise is that *word meanings must be described relative to semantic frames*. FrameNet also considers frames as being partially ordered by several special semantic relations, among which the “inherits” relation stands out as the most informative and restrictive.

Our contention is that Formal Concept Analysis methods and constructions can help greatly in formalizing frames, frame elements and other frame-theoretic concepts. Because of the constructive approach of concept lattices, such formalization would automatically entail visualisation, data-mining and (lattice-oriented) application-building capabilities around the FrameNet data. Our purpose with this paper is to show the abundance of concept lattices in FrameNet and how they can help not only in the above-mentioned roles, but also in bringing formalization to strictly frame-theoretic notions.

In the rest of the paper we first review the general enterprise of FrameNet with special attention to the concept of “coreness” so that the reader may later understand lattices built from a lexicographical perspective. In section 3 we introduce the types of lattices we will be considering and how to build them from FrameNet data. Section 4 is dedicated to the analysis of the *coreness* of frame elements and invokes Formal Concept Analysis to obtain concept lattices that help define the concept of “coresets”, crucial to Frame Semantics, producing some analyses of coreness and coresets as obtained from the FrameNet database. In this section we also pose the question whether frame structure is inherited through the frame hierarchy. We conclude with some suggestions of research into Frame Semantics using concept lattices.

## 2 FrameNet: An Embodiment of Frame Semantics

Regardless of their theoretical interest, lexical resources capable of serving natural language processing applications should at least include the following types of information:

- representations of the meaning of each lexical unit (LU);
- various types of relations between lexical units;
- information about a word’s capacity to combine with other linguistic units to form phrases and sentences;
- semantic information associated with individual words that allows us to interpret the phrases that contain them.