

Using Formal Concept Analysis to Acquire Knowledge about Verbs

Ingrid Falk¹²⁴ Claire Gardent³⁴ Alejandra Lorenzo³⁴

¹INRIA Nancy Grand Est

²Lorraine University, Nancy, France

³CNRS

⁴LORIA

Concept Lattices and their Applications, 2010

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

What this talk is about

An application of Formal Concept Analysis (FCA)
in the domain of Natural Language Processing (NLP)

Starting from lexical resources for French (~ a dictionary)

1. we classify French verbs based on syntactic and semantic features,
 - ▶ \rightsquigarrow using concept lattices.
2. we explore relations/dependencies between syntactic and semantic features, and extend the original lexicon
 - ▶ \rightsquigarrow using association rules.

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs**
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

Verbs in Natural Language Processing Applications

- ▶ NLP applications **analyse** and/or **generate** natural language sentences
- ▶ Verbs are central in natural language sentences . . .
- ▶ Knowledge about their **syntax/semantics** is crucial for NLP applications.

A means to **acquire** and **structure** knowledge about verbs are **verb classifications**.

Verb Classifications

Group together verbs with similar **syntactic** and/or **semantic** behaviour.

Benefits:

On the practical side: capture **generalisations**, reduce the effort of building and maintaining a verb lexicon.

On the theoretical side: verbs belonging to the same class often **share a semantic component**.

Available resources:

For English several large scale resources: FrameNet [Baker et al., 1998], VerbNet [Schuler, 2006] and WordNet [Fellbaum, 1998].

For French restricted in scope (Volem [Saint-Dizier, 1999]) or not sufficiently structured (the LADL tables [Gross, 1975]).

Contributions of this work.

1. we start from a French valency lexicon and use FCA to create a **classification of French verbs**.
2. we show how this classification can be extended to also contain **semantic** information.
3. we apply high confidence association rules to a different lexicon to **extend** the initial lexicon.

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.**
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

The Method

- ▶ build a context from a valency lexicon of French verbs – **Dicovalence**¹ [van den Eynde and Mertens, 2003],
- ▶ compute the lattice – **Galicia**²,
- ▶ filter using concept stability – [Kuznetsov, 2007],
- ▶ compare obtained classification to VerbNet – [Schuler, 2006].

¹<http://bach.arts.kuleuven.be/dicovalence/>

²<http://www.iro.umontreal.ca/galicia/>

Dicovalence [van den Eynde and Mertens, 2003]

- ▶ for French,
- ▶ valency frames for 3936 verbs,
- ▶ created manually.

Example entries

verb	frame and example
<i>manifeste</i>	SUJ:NP, OBJ:NP Cette expression manifeste un dédain réel. This expression shows a real disdain.
<i>manifeste</i>	SUJ:NP, OBJ:NP, A-OBJ Il ne manifeste jamais ses vrais sentiments (à qqn.) He never showed his true feelings (to sb.)

VerbNet [Schuler, 2006]

- ▶ for English,
- ▶ classifies 3626 verbs using 411 classes,
- ▶ created manually,
- ▶ the kind of classification we aim at!

Example class:

Verbs: *batter, beat, bump, butt, drum, hammer, hit, jab, kick, knock, lash, pound, rap, slap, smack, smash, strike, tap*

Frames
SUJ:NP,P-OBJ:PP
SUJ:NP,P-OBJ:PP,P-OBJ:PP
SUJ:NP,OBJ:NP
SUJ:NP,OBJ:NP,P-OBJ:PP
SUJ:NP,DE-OBJ:Ssub

The context:

Objects: the verbs from Dicovalence (eg. *manifestester*),

Attributes: the frames from Dicovalence (eg. SUJ:NP, OBJ:NP)

⇒ a context of 3936 verbs and 136 frames.



A concept lattice with 2115 concepts

Most concepts are not interesting:

- ▶ only 1 or 2 verbs,
- ▶ few frames.

Concept stability

Definition ([Kuznetsov, 2007])

Let (V, F) be a formal concept and $'$ the derivation operator. It's **intensional stability** is defined as:

$$\sigma_i((V, F)) := \frac{|\{A \subseteq V \mid A' = F\}|}{2^{|V|}}$$

- ▶ the proportion of the subsets of the extent which have the same intent.
- ▶ a more stable concept is less dependant on individual members in the extension.

Problem

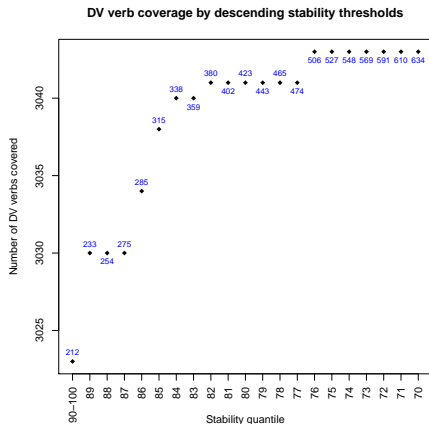
- ▶ keeping only the most stable concepts affects coverage.

Evaluation

Preliminary **quantitative** evaluation.

1. Is the coverage reasonable?
2. Do the classes have good generalisation and factorisation power?

Coverage.



Stability threshold at 86% offers good compromise between:

1. size of frame sets (1-10, 43% of verbs > 2 frames),
2. verb coverage (3038 of 3936, ~ 77%),
3. number of classes (315).

Generalisation and factorisation power.

- Estimation by **comparison with VerbNet**:

Stability threshold	86%	VerbNet
Nb. of classes	315	411
Average class size (verbs)	75.03	14.96
Average class size (frames)	3.51	4.02
Average class size (harmonic mean)	5.98	4.67
Verbs covered	3038/3936/77%	3626

Nbr. of classes, avg. nbr. of verbs per class \rightsquigarrow good generalisation power,
 Harmonic mean of verb set and frame set size \rightsquigarrow good factorisation
 power.

Adding a Semantic Component

- ▶ by adding information from another syntactic-semantic lexical resource for French: **the LADL tables**.

The LADL tables

- ▶ group verbs in tables according to syntactic and/or semantic criteria,
- ▶ developed manually over several years by a large team of linguists,
- ▶ not fine grained enough for NLP.

Example

A table might contain:

- ▶ verbs where the object must be human and the subject is not restricted, or
- ▶ verbs where the object can be a phrase expressing a destination, ...

Method

Lattice is built from:

Objects verbs (as before)

- Attributes**
- ▶ a frame, if the verb has that frame in Dicovallence,
 - ▶ a table identifier, if the verb is in that LADL table.

Filtering

500 most stable concepts, with at least

- ▶ one table identifier,
- ▶ two valency frames.

Results

- ▶ 2 – 6 valency frames,
- ▶ 9 – 237 verbs,
- ▶ coverage: 62% of intersection (2192 of 3536).

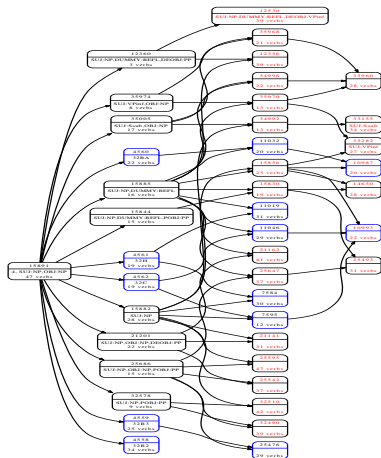
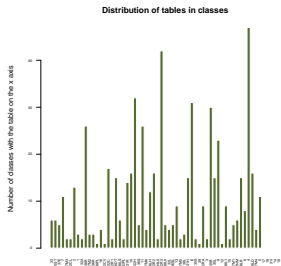


Figure 4 Classes with table 4

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.**
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

- ▶ **Association rules** $A \rightarrow B$ relate sets of attributes.
- ▶ in our case they describe **dependencies between frames**.

We extract rules from the context obtained from Dicovalence ...

And address two questions:

1. How interesting are the dependencies inferred from Dicovalence?
2. How can we extend Dicovalence using these rules.

How interesting are rules obtained from Dicovalence?

Extracted rules:

→ minimal non redundant association rules³, i.e. rules $F_1 \rightarrow F_2$ s.t. F_2 closed itemset and F_1 is minimal generator of F_2 .

Metrics to assess “interestingness”:

Confidence: the probability to have F_2 given F_1 ,

Lift > 1: antecedent and consequent appear more often together than expected.

For our data:

- ▶ confidence of most rules between 98% and 100%
- ▶ lift > 1 for almost all rules.

⇒ Most rules are interesting.

Method:

- ▶ select a set of *good* rules, ...
- ▶ apply these rules to the frames of (verb, frame) pairs given by LADL,
- ▶ add the inferred set of (verb, frame) pairs to Dicovalence.

A rule is *good* if

- ▶ it is applicable to a large proportion of LADL (verb, frame) pairs (**applicability**),
- ▶ when it's applicable it gives new information, ie. the involved frames are in the rule's antecedent and the frames in the consequent are not in LADL (**usefulness**).

Applying association rules I

Metrics: **applicability** and **usefulness**.

For a rule r :

Applicability: proportion of LADL verbs where r can be applied.

Let

- ▶ V_{LADL}^r the set of verbs in LADL which have a frame in the antecedent of r (where r can be applied) and
- ▶ V_{LADL} the set of LADL verbs.

$$applicability(r) := \frac{|V_{LADL}^r|}{|V_{LADL}|}$$

Applying association rules II

Usefulness: proportion of inferred frames not in LADL.

For a rule r let

- ▶ F_{LADL}^r set of frames in LADL for which there is a verb $\in V_{LADL}^r$ and
- ▶ $NewF_{LADL}^r$ frames in the consequent of F_{LADL}^r and not present in LADL.

$$usefulness(r) = \frac{|NewF_{LADL}^r|}{|F_{LADL}^r|}$$

ie. the proportion of discovered frames and the number of frames contained in the verb entries to which the rule applies.

Applying association rules III

Selecting 30 rules with highest applicability:

- ▶ most are applicable to $< 5\%$ of LADL verbs,
- ▶ most are *useful*: they permit to infer a large number of new LADL frames (usefulness is 10%-40%),
- ▶ they are reliable:
 - ▶ good confidence: 0.72 – 1,
 - ▶ lift > 1

Applying these rules to (verb, frame) pairs from LADL:

- ▶ \rightsquigarrow 1435 inferred (verb,frame) pairs,
- ▶ more than when using support, lift and confidence for selecting the rules (1157).

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

Conclusion

Results suggest that FCA is an appropriate framework for identifying verb classes grouping simultaneously verbs and associated syntactic and semantic information:

- ▶ concepts obtained by FCA naturally model the verb/frame association,
- ▶ FCA permits soft clustering – a data element may belong to several classes,
- ▶ stable concepts have good generalisation and factorisation power and linguistically sound empirical content.
- ▶ Association rules naturally capture **alternations**, ie. frames that are often true simultaneously.
- ▶ Metrics in the literature permit ranking them according to different criteria . . .
- ▶ and association rules can be used to extend a lexicon.

Future Work




- ▶ Use association rules to extend the data set.
- ▶ Improve the classification
 - ▶ use more and more detailed syntactic and semantic features,
 - ▶ model and make use of relations and dependencies between attributes.
- ▶ Use classification to infer more semantic information (θ -grids for semantic role labeling).

Thank you!

Summary

- 1 Overview
- 2 Motivation: NLP and Verbs
- 3 Acquiring Verb Classes with FCA.
 - Lexical Resources
 - The concept lattice.
 - Filtering
 - Evaluation
 - Adding a Semantic Component.
- 4 Using association rules.
 - Extending Dicovalence
- 5 Conclusion and future work.
- 6 References

Lexical Resources. I

-  Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998).
The berkeley FrameNet project.
International Conference on Computational Linguistics, Montreal,
Quebec, Canada.
-  Fellbaum, C., editor (1998).
WordNet: An Electronic Lexical Database.
MIT Press, Cambridge, MA.
-  Gross, M. (1975).
Méthodes en syntaxe.
Hermann, Paris.

Lexical Resources. II



Levin, B. (1993).

English Verb Classes and Alternations: a preliminary investigation.

University of Chicago Press, Chicago and London.



Saint-Dizier, P. (1999).

Alternation and verb semantic classes for french: Analysis and class formation.

In *Predicative forms in natural language and in lexical knowledge bases*. Kluwer Academic Publishers.




Schuler, K. K. (2006).

VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon.

PhD thesis, University of Pennsylvania.

Lexical Resources. III

 van den Eynde, K. and Mertens, P. (2003).

La valence: l'approche pronominale et son application au lexique verbal.

Journal of French Language Studies

FCA



Kuznetsov, S. O. (2007).

On stability of a formal concept.

Annals of Mathematics and Artificial Intelligence, 49(1-4):101–115.