

Exploring the Variability of Interconnected Product Families with Relational Concept Analysis

Jessie Carbonnel Marianne Huchard Clémentine Nebut

IRD (Institut de Recherche pour le Développement), Montpellier
LIRMM, CNRS and Université de Montpellier

23rd International Systems and Software Product Line Conference
Sept 9-13 Paris, France

Outline

- 1 Introduction
- 2 Interconnected product descriptions and their representation
- 3 Variability extraction with FCA/RCA
- 4 FCA/RCA Benefits
- 5 Conclusion

Context

Large or complex product lines

- Multi-stakeholder
- Multi-viewpoint
- ▶ **Multi-product**
- Multi-stage

▶ **Variability extraction and representation**

- **Understanding**
- **Migration**
- **Refactoring**
- **Adaptation**
- **Recommendation**

Objective

Targeted description elements

- Features
- Multi-valued attributes
- Cardinalities (for feature or feature group)
- ▶ References (for cross-product lines)

Targeted variability relationships extraction

- ▶ Binary implications, co-occurrences
- ▶ Mutual exclusions
- ▶ Groups (OR, XOR)
- Constraints, Logic formulas
- ▶ ▶▶▶ *Intermediate structures*

Method

Using a framework

- In correspondence with propositional/predicate/description logics background
- Sound and complete extraction capabilities
- Canonical and exhaustive constructions
- Graphical representation capabilities
- Reusable, Extensible

Underlying theory

- Formal Concept Analysis
 - ▶ Configurations + Features
- Relational Concept Analysis
 - ▶ *Configurations + Features*
 - + Configuration categories
 - + relationships between configurations of different categories

Method

Approach in a nutshell

- Apply FCA/RCA
- Analyze and interpret the conceptual structures
 - ▶ Mapping to variability relationships (e.g. for refactoring)
 - ▶ Exploring the conceptual structures (e.g. for recommendation)

Outline

- 1 Introduction
- 2 Interconnected product descriptions and their representation**
- 3 Variability extraction with FCA/RCA
- 4 FCA/RCA Benefits
- 5 Conclusion

A typical example

Interconnected Product Comparison Matrices (PCM)

- Several yes/no columns, column headers are described in another PCM
- One multi-valued column, values are rows in another PCM

Comparison of accounting software

Package	License	Windows	Mac OS	Linux	Market focus	GL	Language
GNU Enterprise	GPL	Yes	Yes	Yes	Mid to high end market		Python
Odoo (CE)	LGPL3	Yes	Yes	Yes	All markets	Yes	Python
Tryton	GPL	Yes	Yes	Yes	Low to mid-market	Yes	Python
Dolibarr	GPL	Yes	Yes	Yes	Low to mid-market	Yes	PHP
LedgerSMB	GPL	Yes	Yes	Yes	Mid-market	Yes	Perl
SQL-	GPL	Yes	Yes	Yes	Mid-	Yes	Perl

Comparison of operating systems

GNU/Linux	Notable contributors include: Richard Stallman for GNU Project and Linux Torvalds for Linux and the Unixes they emulated; Red Hat, Debian Project See: Comparison of Linux distributions and Linux Kernel#Development	1 Cl
Haiku	Haiku Inc.	
macOS	Apple Inc.	
macOS Server	Apple Inc.	

Comparison of free and open-source software licenses

License	Author	Latest version	Publication date
Academic Free License ^[10]	Lawrence E. Rosen	3.0	2002
Affero General Public License	Affero Inc	2.0	2007
Apache License	Apache Software	2.0	2004

Cross-family variability, Multi Product Lines

- Survey on MPL [Holl et al., 2012]
- Feature Models Reference [Czarnecki et al., 2002, Czarnecki et al., 2004]
- Orthogonal Variability Model [Pohl et al., 2005]
- Embedded System Families [Streitferdt et al., 2005]
- PLs that Supply Other PLs [Trujillo et al., 2007]
- Modular Embedded Software Product Lines [Friess et al., 2007]
- Dependent software product lines [Rosenmüller et al., 2008]
- Modular Feature Models [Bagheri et al., 2011]
- Invar, Bridging Heterogeneous Variab. Model. App. [Dhungana et al., 2011]
- pure::variants ^a
- Gears ^b
- ... (non exhaustive list)

^a<http://www.pure-systems.com>

^b<http://www.biglever.com>

Outline

- 1 Introduction
- 2 Interconnected product descriptions and their representation
- 3 Variability extraction with FCA/RCA**
- 4 FCA/RCA Benefits
- 5 Conclusion

Formal Concept Analysis

Principles [Ganter and Wille, 1999]

- General theory on (Galois) connections between lattices structures [Birkhoff, 1940, Barbut and Monjardet, 1970]
- Restricted in FCA to **Configurations + Boolean Features** = Concepts + Hierarchy

	document type	typesetting quality
OfficeSuite		
MsOff	spread sheet, text vector graphics, presentation	standard
libreOff	spread sheet, text vector graphics, presentation	standard
TexM	text, math formula, presentation	high
TexLxM	text, math formula, presentation	high
ScWrd	text, math formula, presentation	high
Scribs	layout design, vector graphics	standard
Indesig	layout design, vector graphics	standard

product description table

	spreadSheet	text	formula	layoutDesign	vectorGraphics	presentation	qualityTS	StandardTS	HighTS
OfficeSuite									
MsOff	x	x			x	x	x	x	
libreOff	x	x			x	x	x	x	
TexM		x	x			x	x		x
TexLxM		x	x			x	x		x
ScWrd		x	x			x	x		x
Scribs				x	x		x	x	
Indesig				x	x		x	x	

binary table Configurations +
Boolean Features (formal context)

Simplified description of dummy office suites

FCA

Principles

Configurations + Features = **Concepts** + Hierarchy

	spreadSheet	text	formula	layoutDesign	vectorGraphics	presentation	quality TS	StandardTS	HighTS
OfficeSuite									
MsOff	x	x			x	x	x	x	
libreOff	x	x			x	x	x	x	
TexM		x	x			x	x		x
TexLxM		x	x			x	x		x
ScWrd		x	x			x	x		x
Scribs				x	x		x	x	
Indesig				x	x		x	x	

formal context

Concept_OfficeSuite_5
text
presentation
quality TS
MsOff
libreOff
TexM
TexLxM
ScWrd

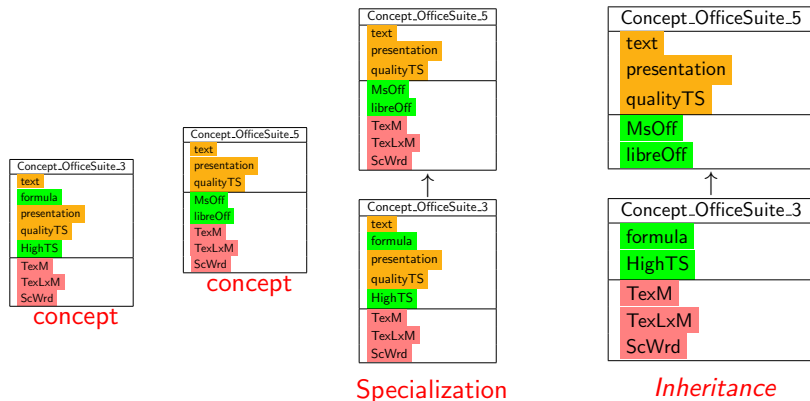
concept

Concept = maximal configuration set sharing maximal feature set

FCA

Principles

Configurations + Features = Concepts + **Hierarchy**

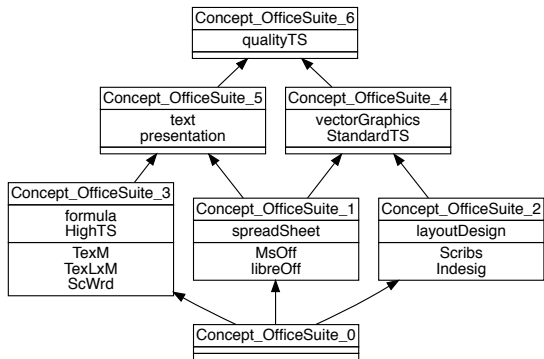


Features: top-down
Configurations: bottom-up

FCA

Principles

Configurations + Features = Concepts + **Hierarchy**

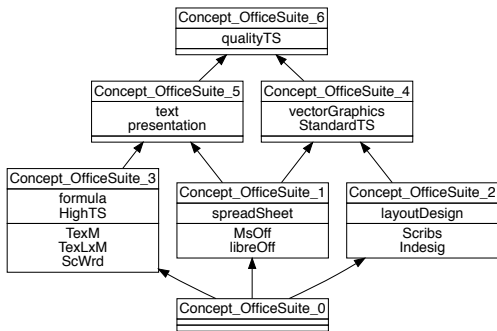


Concept lattice

FCA

Variability extraction [Loesch and Ploedereder, 2007, Ryssel et al., 2011]

Implications and co-occurrences



Concept_OfficeSuite_3
subconcept of
Concept_OfficeSuite_5

formula \Rightarrow text

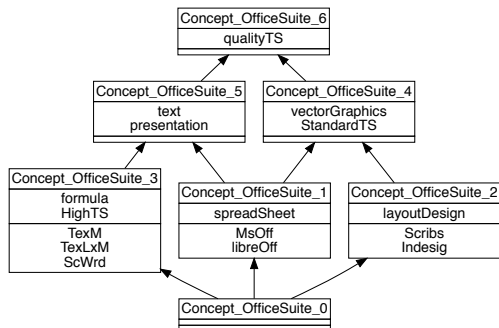
Both introduced in
Concept_OfficeSuite_5

text \Leftrightarrow presentation

FCA

Variability extraction

Mutex



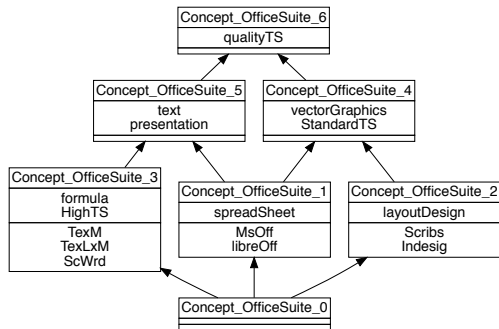
Concept_OfficeSuite_3
and
Concept_OfficeSuite_2
have *no common subconcept*
introducing a configuration

formula $\Rightarrow \neg$ layoutDesign

FCA

Variability extraction

Xor Groups



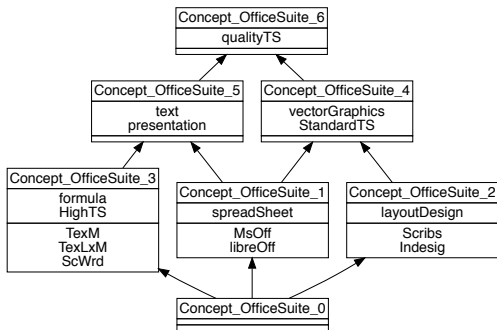
*Configurations of
Concept_OfficeSuite_3
and
Concept_OfficeSuite_4
form a **partition** of
the configurations of
Concept_OfficeSuite_6*

qualityTS
Xor
{ HighTS, StandardTS }

FCA

Variability extraction

Candidate Or Groups



*Configurations of
Concept_OfficeSuite_6
are covered by
non disjoint union
of these of
Concept_OfficeSuite_5
and
Concept_OfficeSuite_4*

Office suite

Or

{ text, vector graphics }

RCA [Hacene et al., 2013]

Principles

Configurations + Features + **Configuration categories** + relationships between configurations of different categories = Concepts in Inter-connected Hierarchies

Config	config	computer	desktop	labtop	windows	macOS	linux
Secr1	x	x	x		x		
Secr2	x	x	x		x		
Com1	x	x	x		x		
Com2	x	x	x		x		
CEO1	x	x		x		x	
CEO2	x	x		x		x	
ITE1	x	x		x			x
ITE2	x	x		x		x	
ITE3	x	x		x		x	

	spreadSheet	text	formula	layoutDesign	vectorGraphics	presentation	qualityTS	StandardTS	HighTS
OfficeSuite									
MsOff	x	x			x	x	x	x	
libreOff	x	x			x	x	x	x	
TexM		x	x			x	x		x
TexLxM		x	x			x	x		x
ScWrd		x	x			x	x		x
Scribs				x	x				
Indesig				x	x				

	ticket	IDE	modelling	bdAdmin
DevTool				
Jiraf	x			
Bugzz	x			
Intelli		x		
Eclip		x	x	
Modellioo			x	
Rsaa			x	
PHHmiadm				x
MiSeqwb				x

	mail	videoConf	chat
CollabTool			
Outlak	x		
Zimbro	x		
BlueBird	x		
Skyte		x	x
ApachOM		x	x

	planning	budgetChart	Gantt	Pert	analytics	sales	time	resource
MngtSoft								
eclipPPM	x		x	x	x	x	x	x
ganttPrjct	x		x					
Mansit	x		x					
orPrimaPPM	x	x	x	x	x	x	x	x
ZhooOS	x	x	x	x	x	x	x	x

RCA

Principles

Configuration categories + Configurations + Features + **relationships between configurations of different categories** = Concepts in Inter-connected Hierarchies

OffSuit	MsOff	libreOff	TexM	TexLxM	ScWrd	Scribs	Indesig
Secr1	x						
Secr2	x						
Com1	x					x	
Com2	x						x
CEO1	x						
CEO2		x					
ITE1		x	x				
ITE2	x			x	x		
ITE3		x		x			

devTool	Jiraf	Bugzz	Intelli	Eclip	Modelico	Rsa	PHmiadm	MISeqwb
Secr1								
Secr2								
Com1								
Com2								
CEO1								
CEO2								
ITE1		x	x					x
ITE2	x			x	x			
ITE3	x			x	x		x	

collabTool	Outlak	Zimbro	BlueBird	Skyte	ApachOM
Secr1		x			
Secr2		x			
Com1		x			
Com2		x			
CEO1	x			x	
CEO2		x		x	
ITE1			x		x
ITE2		x		x	
ITE3		x		x	

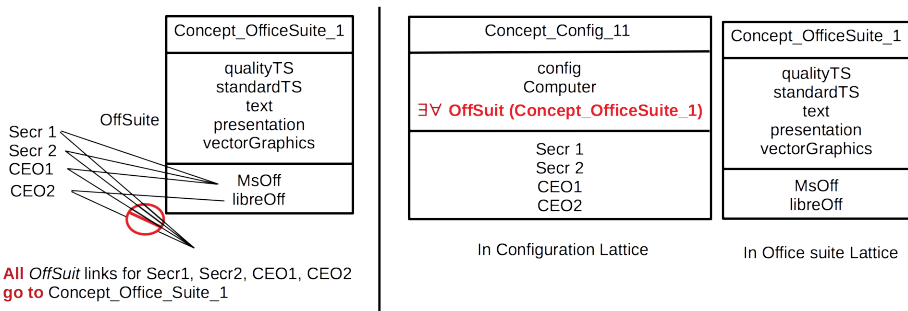
mngtSoft	eclipPPM	gantPrjct	Mansit	orPrimaPPM	ZhoosOS
Secr1					
Secr2					
Com1					
Com2					
CEO1				x	
CEO2	x				
ITE1					
ITE2					
ITE3					

RCA

Configuration categories + Configurations + Features + relationships between configurations of different categories = **Concepts in Inter-connected Hierarchies**

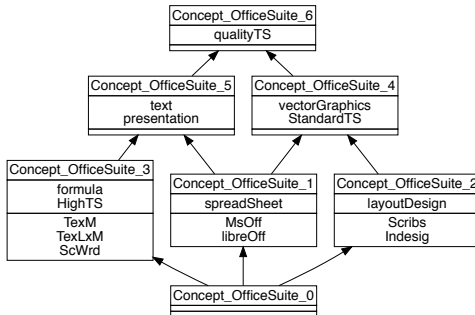
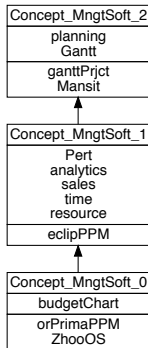
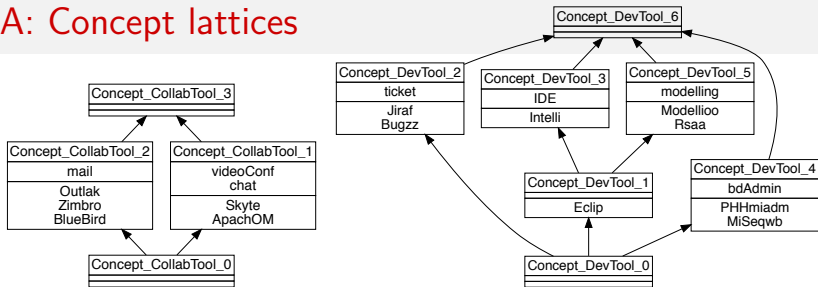
Introducing relational attributes

a quantifier	a relation	a concept
$\exists \forall$ All links	<i>OffSuit</i> for <i>OffSuit</i> go to	<i>(Concept_OfficeSuite_1)</i> <i>(Concept_OfficeSuite_1)</i> extent ¹

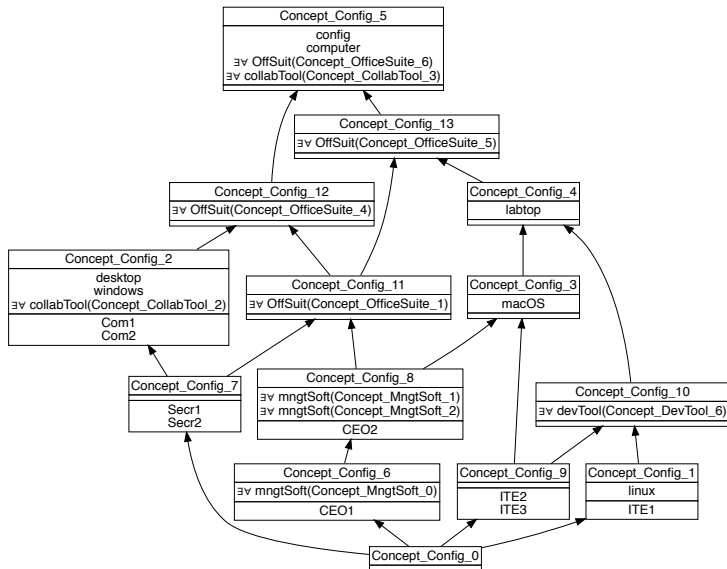


¹and to super concepts by construction

RCA: Concept lattices



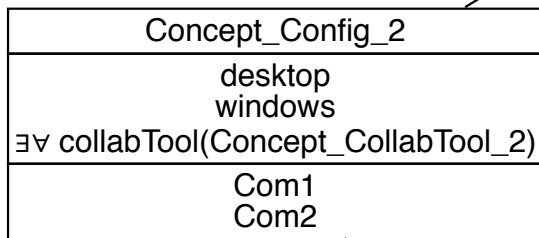
RCA: (User) Configuration Concept lattice



Variability extraction with RCA

An example of co-occurrence

desktop $\Leftrightarrow \exists \forall \text{collabTool}(\text{Concept_CollabTool_2})$ —includes mail

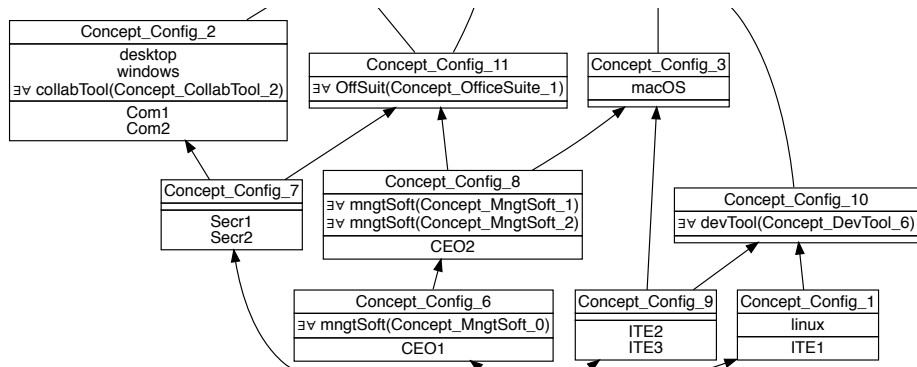


Variability extraction with RCA

An example of Candidate Xor

Concept_Config_2 (Secr. and Com. employees), *Concept_Config_8* (CEOs), *Concept_Config_10* (ITEs) partition all user configurations

User configuration = *Xor* { Secr/Com, CEOs, ITEs }



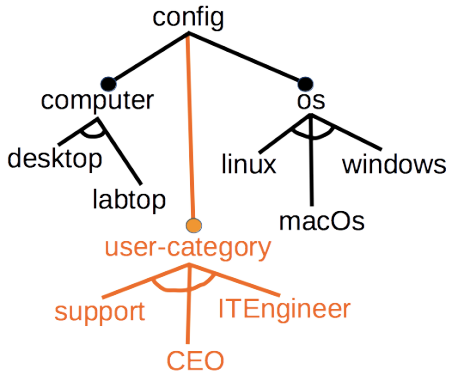
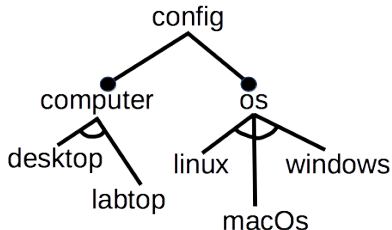
Outline

- 1 Introduction
- 2 Interconnected product descriptions and their representation
- 3 Variability extraction with FCA/RCA
- 4 FCA/RCA Benefits**
- 5 Conclusion

Interconnected FMs synthesis

Suggestions for Feature Model Synthesis

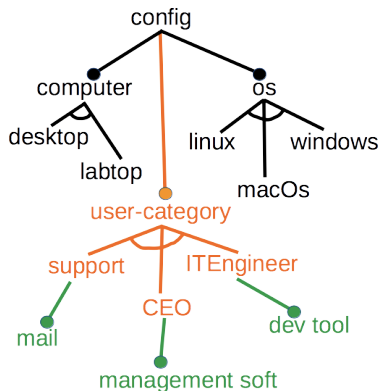
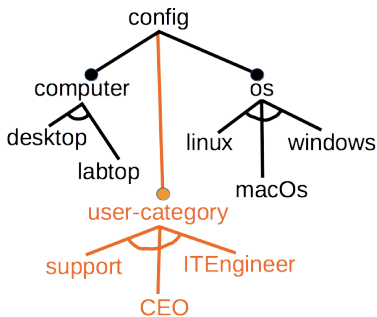
Using *Concept_Config_2* (Secr. and Com. employees), *Concept_Config_8* (CEOs), *Concept_Config_10* (ITEs)



Interconnected FMs synthesis

Suggestions for Feature Model Synthesis

Using relational attributes $\exists \forall collabTool(Concept_CollabTool_2)$,
 $\exists \forall mngtSoft(Concept_MngtSoft_2)$, $\exists \forall devTool(Concept_DevTool_6)$

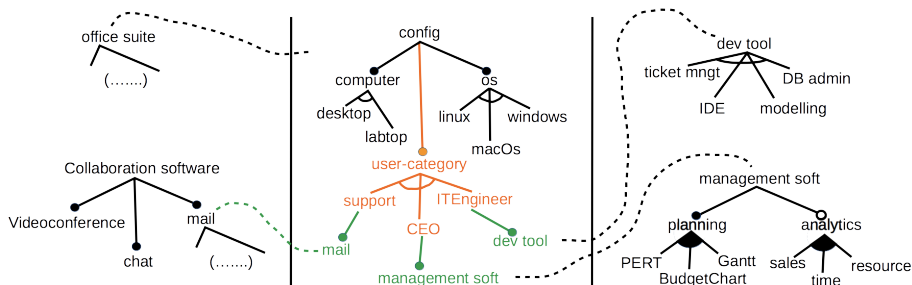


Interconnected FMs synthesis

Suggestions for Feature Model Synthesis

Using relational attributes

- $\exists \forall \text{collabTool}(\text{Concept_CollabTool_2})$ internal lattice node \rightarrow internal FM node,
- $\exists \forall \text{mngtSoft}(\text{Concept_MngtSoft_2})$ lattice top node \rightarrow FM root,
- $\exists \forall \text{devTool}(\text{Concept_DevTool_6})$ lattice top node \rightarrow FM root

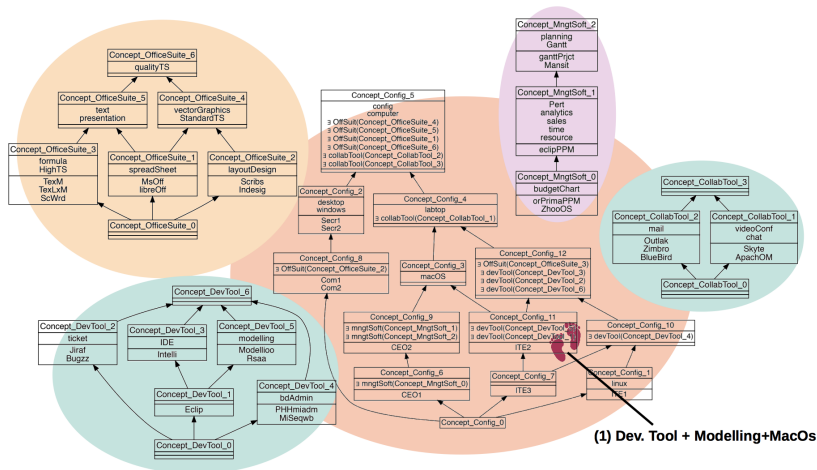


Recommendation / Exploration

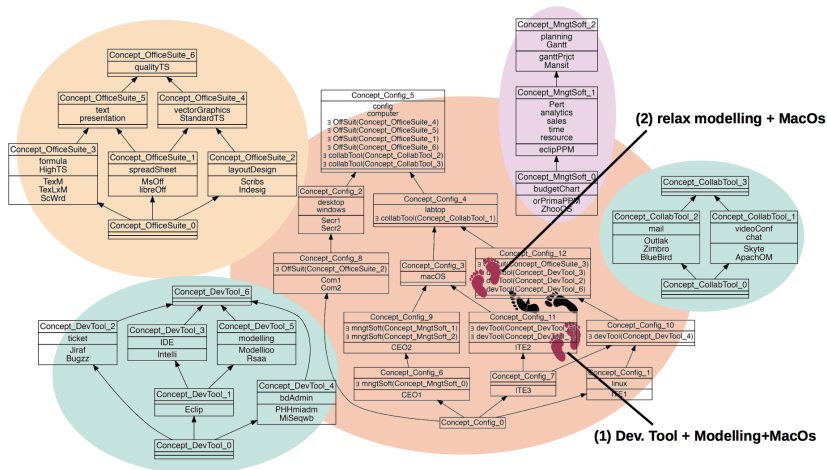
Using the conceptual structures, which embed variability information

- Exploring one structure
- Jumping from one structure to another
- Selecting categories

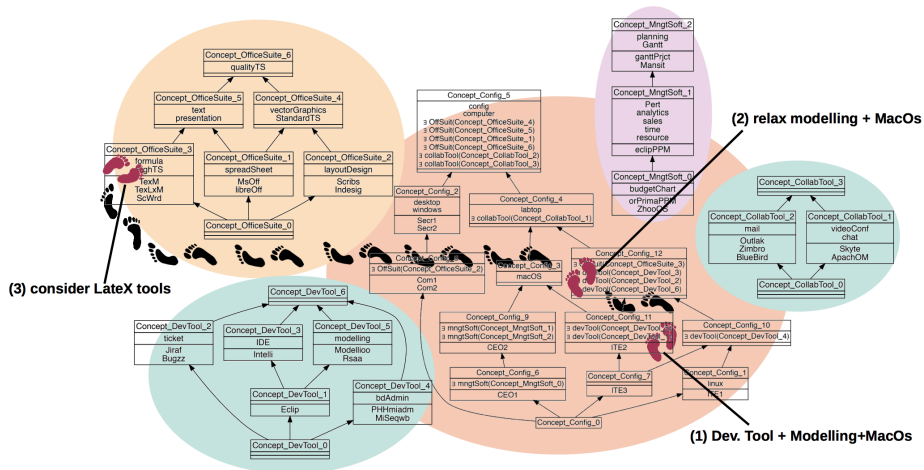
Recommendation / Exploration



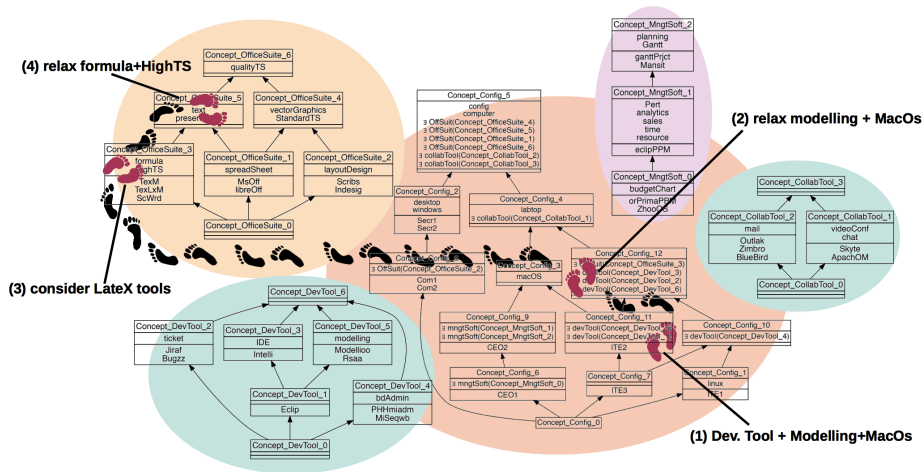
Recommendation / Exploration



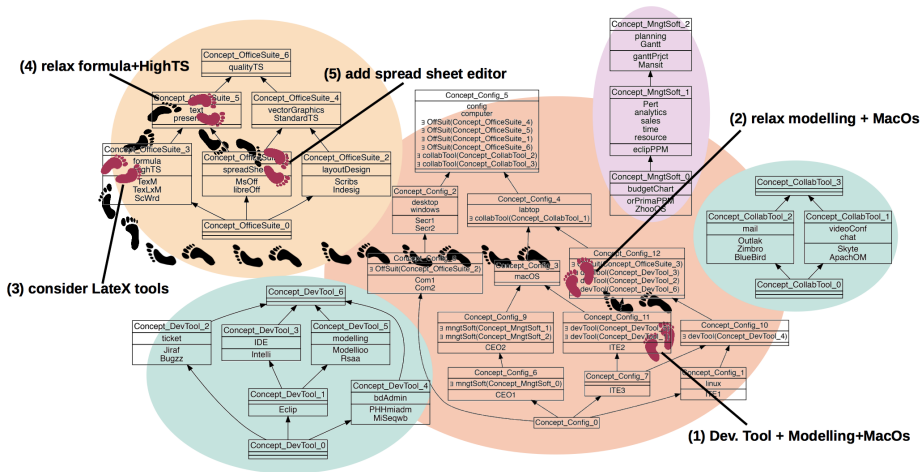
Recommendation / Exploration



Recommendation / Exploration



Recommendation / Exploration



Outline

- 1 Introduction
- 2 Interconnected product descriptions and their representation
- 3 Variability extraction with FCA/RCA
- 4 FCA/RCA Benefits
- 5 Conclusion

Conclusion

Variability relationship extraction

- In interconnected product families
- Assisted by Relational Concept Analysis
- Introduction of cross-family features and variability relationships

Exploitation

- Assistance to build interconnected variability representations, e.g. FMs
- Exploration of a set of interconnected configurations to find or create complex configurations

Perspectives

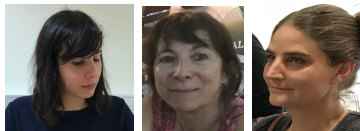
Deepen methodological aspects

- Assistance in reading/traversing the interconnected conceptual structures
- Exploit the variety of scaling quantifiers: \exists , $\exists\forall$, $\exists \supseteq$, $\exists\forall_{\geq n\%}$, $\exists \supseteq_{\geq n\%}$
- Define a complete assistance method for building interconnected FMs synthesis (and suggest new connexion forms between the FMs)
- Exploration of large configuration offerings through local construction of conceptual structures

Validation

- PCMs
- Sampling results, e.g. JHipster sampling [Halin et al., 2017]
- Listening proposals!

Thank you!





Bagheri, E., Ensan, F., Gasevic, D., and Boskovic, M. (2011).
Modular feature models: Representation and configuration.
Journal of Research and Practice in Information Technology, 43(2):109–140.



Barbut, M. and Monjardet, B. (1970).
Ordre et classification – Algèbre et combinatoire, Tome 2.
Hachette, Paris.



Birkhoff, G. (1940).
Lattice Theory.
Colloquium publications. American Mathematical Society.



Czarnecki, K., Bednasch, T., Unger, P., and Eisenecker, U. W. (2002).
Generative Programming for Embedded Software: An Industrial Experience Report.
In *Proceedings of the 1st ACM SIGPLAN/SIGSOFT Conference on Generative Programming and Component Engineering (GPCE'02)*, Lecture Notes in Computer Science, pages 156–172. Springer.



Czarnecki, K., Helsen, S., and Eisenecker, U. W. (2004).
Staged Configuration Using Feature Models.
In *Proceedings of the 3rd International Conference on Software Product Lines (SPLC'04)*, Lecture Notes in Computer Science, pages 266–283. Springer.



Dhungana, D., Seichter, D., Botterweck, G., Rabiser, R., Grünbacher, P., Benavides, D., and Galindo, J. A. (2011).
Configuration of multi product lines by bridging heterogeneous variability modeling approaches.
In *Software Product Lines - 15th International Conference, SPLC 2011, Munich, Germany, August 22-26, 2011*, pages 120–129.



Friess, W., Sincero, J., and Schroeder-Preikschat, W. (2007).
Modelling Compositions of Modular Embedded Software Product Lines.
In *25th Conference on IASTED International Multi-Conference*, pages 224–228. ACTA Press.



Ganter, B. and Wille, R. (1999).
Formal concept analysis - mathematical foundations.
Springer.



Hacene, M. R., Huchard, M., Napoli, A., and Valtchev, P. (2013).
Relational concept analysis: mining concept lattices from multi-relational data.
Ann. Math. Artif. Intell., 67(1):81–108.



Halin, A., Nuttinck, A., Acher, M., Devroey, X., Perrouin, G., and Heymans, P. (2017).

Yo variability! jhipster: a playground for web-apps analyses.

In *Proceedings of the 11th International Workshop on Variability Modelling of Software-intensive Systems (VaMoS'17)*, pages 44–51. ACM.



Holl, G., Grünbacher, P., and Rabiser, R. (2012).

A systematic review and an expert survey on capabilities supporting multi product lines.

Information and Software Technology, 54(8):828 – 852.

Special Issue: Voice of the Editorial Board.



Loesch, F. and Ploedereder, E. (2007).

Restructuring Variability in Software Product Lines using Concept Analysis of Product Configurations.

In *Proceedings of the 11th European Conference on Software Maintenance and Reengineering, Software Evolution in Complex Software Intensive Systems (CSMR'07)*, pages 159–170. IEEE Computer Society.



Pohl, K., Böckle, G., and van der Linden, F. J. (2005).

Software Product Line Engineering: Foundations, Principles, and Techniques.

Springer Science & Business Media.



Rosenmüller, M., Siegmund, N., Kästner, C., and ur Rahman, S. S. (2008).

Modeling dependent software product lines.

In *Proceedings of the GPCE Workshop on Modularization, Composition and Generative Techniques for Product Line Engineering (McGPLE'08)*, pages 13–18.



Ryssel, U., Ploennigs, J., and Kabitzsch, K. (2011).

Extraction of feature models from formal contexts.

In *Workshop Proceedings (Volume 2) of the 15th International Conference on Software Product Lines (SPLC'11)*, pages 4:1–4:8. IEEE Computer Society.



Streitferdt, D., Sochos, P., and Philippow, C. H. I. (2005).

Configuring Embedded System Families Using Feature Models.

In *Net.ObjectDays*, pages 339–350. Gesellschaft für Informatik.



Trujillo, S., Kästner, C., and Apel, S. (2007).

Product Lines that Supply Other Product Lines: A Service-Oriented Approach.

In *SPLC Workshop: Service-Oriented Architectures and Product Lines - What is the Connection?*

Recommendation / Exploration

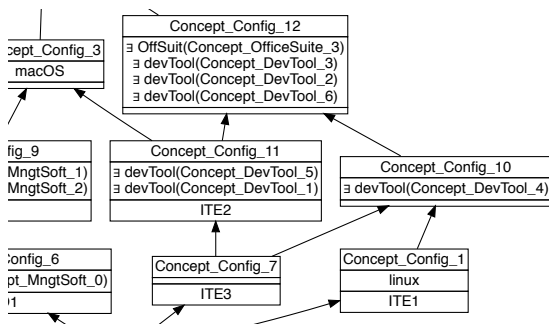
By navigation in the user configuration lattice (quantifier \exists)

Looking for modelling (CDevT5) and macOS (in CConf3) \rightarrow CConf11 (ITE2, ITE1)

From CConf11, adding DBAdmin (CDevT4) inherited from CConf10 \downarrow CConf7

Or, from CConf11, Relaxing modelling needs (CDevT5, CDevT1) \uparrow CConf12

From CConf12 \rightarrow COffSuite3 \rightarrow ... (the story continues in office suite lattice)



Recommendation / Exploration

By navigation in the user configuration lattice

From COffSuite3 → a Latex tool can be selected, or a specific scientific formula editor

If we relax formula and high quality typesetting ↑ COffSuite5

From COffSuite5: adding spread sheet editor ↓ COffSuite1 → select MsOff or libreOff

This creates new user configurations, possibly verifying different constraints

