



---

# Combining Implication Rules and Conceptual Analysis to Learn from a Pesticidal Plant Knowledge Base

---

**Lina Mahrach**, Alain Gutierrez, Marianne Huchard  
Priscilla Keip, Pascal Marnotte, Pierre Silvie and Pierre Martin

ICCS'2021  
Virtual Bolzano, France  
September, 21, 2021

AIDA, LIRMM, PHIM Plant Health Institute

- ❖ Collect data from scientific publications on the uses (45.000 on Sept. 2021) of pesticidal plants to identify alternatives to chemical pesticides and antibiotics

Base de données

Plante		Informations plante (caractéristiques)	Ravageurs, maladies ou auxiliaires		Code de classification	Source de l'information			
Nom latin			Organisme à protéger (espèce de plante)	Nom latin		Auteur	Année	Titre article ou du document	Revue
<i>Citrus sinensis</i>			<i>Manodora myristica</i>	<i>Aspergillus flavus</i>		100	Djeuga	2017 Morphological and Internatic	
<i>Citrus sinensis</i>			<i>Manodora myristica</i>	<i>Fusarium oxysporum</i>		40	Djeuga	2017 Morphological and Internatic	
<i>Citrus sinensis</i>			<i>Manodora myristica</i>	<i>Fusarium oxysporum</i>		60	Djeuga	2017 Morphological and Internatic	
<i>Citrus sinensis</i>			<i>Manodora myristica</i>	<i>Fusarium oxysporum</i>		80	Djeuga	2017 Morphological and Internatic	
<i>Citrus sinensis</i>			<i>Manodora myristica</i>	<i>Fusarium oxysporum</i>		100	Djeuga	2017 Morphological and Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		99	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		78	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		150	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		33	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		62	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		12	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		25	Dakole	2016 Antifungal potentia Internatic	
<i>Callistemon citrinus</i>			<i>Lycopersicum esculen</i>	<i>Phytophthora infestans</i>		100	Dakole	2016 Antifungal potentia Internatic	

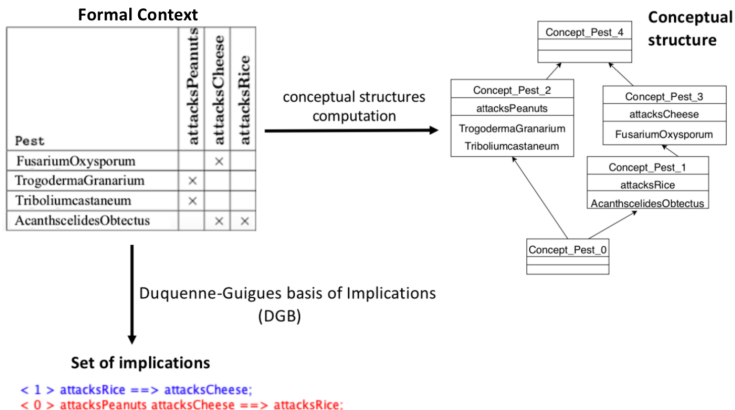
- ❖ Objective: Extract knowledge formulated as implication rules

Silvie P., Martin P., Huchard M., Keip P., Gutierrez A., Sarter S., 2021. Prototyping a knowledge-based system to identify botanical extracts for plant health in Sub-Saharan Africa. *Plants* 10 (5):896-920.  
<https://doi.org/10.3390/plants10050896>

- ❖ Background
- ❖ Dealing with data structured in one table
- ❖ Data structured in several tables
- ❖ Methodology based on Relational Concept Analysis (RCA)
- ❖ Results
- ❖ Discussion
- ❖ Conclusion and future work

# Background

Formal Concept Analysis [Ganter and Wille, 1999]



Guigues, J. L. ; Duquenne, V.. Familles minimales d'implications informatives résultant d'un tableau de données binaires. Mathématiques et Sciences humaines, Tome 95 (1986) , pp. 5-18.

### Formal context and Galois connection "

- ❖  $K = (G, M, I)$   $G$  object set  $M$  attribute set  $I \subseteq G \times M$
- ❖ For  $O \subseteq G$ , shared attributes  $O' = \{m \mid \forall g \in O, (g, m) \in I\}$
- ❖ For  $A \subseteq M$ , sharing objects  $A' = \{g \mid \forall m \in A, (g, m) \in I\}$

### Implication

- ❖  $A \implies B$  with  $A' \subseteq B'$
- ❖ Scope =  $|A'|$
- ❖ Support =  $|A'|/|G|$ .

### Pseudo-intent

Attribute set  $P_i \subseteq M$  s.t:

- ❖  $P_i$  is not an intent ( $P_i'' \neq P_i$ );
- ❖ For any other pseudo-intent  $P_j \subset P_i$ ,  $P_j'' \subset P_i$ .

### Duquenne-Guigues basis of Implications (Duquenne and Guigues, 1986)

- ❖ Implication set  $\{P_i \implies P_i'' \mid P_i \text{ is a pseudo-intent}\}$
- ❖ Cardinality minimal basis of non redundant implications

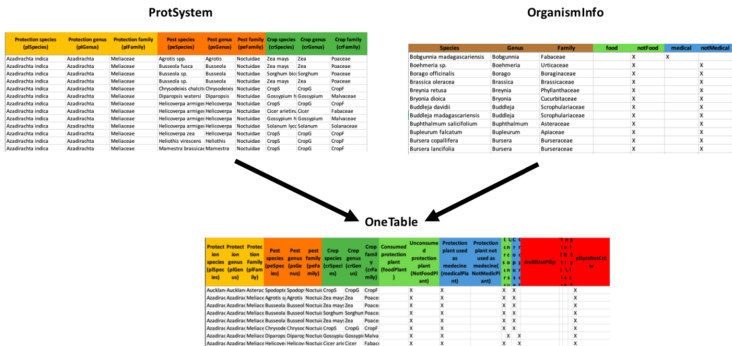
### Computation

- ❖ Lincbo (Janostik et al., 2021)
- ❖ FCA4J (<http://www.lirmm.fr/fca4j>)

# Dealing with data structured in one table

## Knomana: Knowledge management on pesticides plants in Africa

### Merging 2 tables in one table



New **complex** attributes were introduced, e.g. **MultiUsePIGe\_** which refers to a genus of plant that can be used as a pesticidal plant and as a protected crop in separate uses

# Dealing with data structured in one table

## Knomana: Knowledge management on pesticides plants in Africa

OneTable

Protect	Protect	Protect	Pest	Pest	Pest	Crop	Crop	Crop	Consumed	Unconsumed	Protection	Protection						
ion	ion	ion	genus	genus	genus	species	species	family	protection	protection	as	not used as						
(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec	(pISpec						
us)	us)	us)	us)	us)	us)	us)	us)	us)	us)	us)	us)	us)						
Aucklan	Aucklan	Asterac	Spodopt	Spodopt	Noctui	Crop5	CropG	CropF	X	X			X	X				X
Azadrac	Azadrac	Mellace	Agrotis	Agrotis	Noctui	Zea mayi	Zea	Poace	X	X			X	X				X
Azadrac	Azadrac	Mellace	Busseola	Busseola	Noctui	Zea mayi	Zea	Poace	X	X			X	X				X
Azadrac	Azadrac	Mellace	Busseola	Busseola	Noctui	Sorghum	Sorghum	Poace	X	X			X	X				X
Azadrac	Azadrac	Mellace	Busseola	Busseola	Noctui	Zea mayi	Zea	Poace	X	X			X	X				X
Azadrac	Azadrac	Mellace	Chrysode	Chrysode	Noctui	Crop5	CropG	CropF	X	X			X	X				X
Azadrac	Azadrac	Mellace	Diparop	Diparop	Noctui	Gossypii	Gossypii	Malva	X	X			X	X				X
Azadrac	Azadrac	Mellace	Helicov	Helicov	Noctui	Cicer aris	Cicer	Fabac	X	X			X	X				X



DGB

*PeFamily\_Noctuidae, MedicalCrop\_X, FoodCrop\_* ⇒ *PeGenus\_Spodoptera, FoodPlant\_*,  
*MedicalPlant\_*, *MultiUsePlSp\_*, *MultiUsePlGe\_*, *MultiUsePlFa\_*, *PlSpIsCrEw\_*, *PlGeIsCrEw\_*,  
*PlFamily\_Asteraceae, PeSpecies\_SpodopteraLitura, CrSpIsPlEw\_X*,  
*CrGeIsPlEw\_X, PlSpecies\_WollastoniaDentata, PlGenus\_Wollastonia*,  
*CrSpecies\_RicinusCommunis, CrGenus\_Ricinus, CrFamily\_Euphorbiaceae (OT1)*

*PeGenus\_Spodoptera, PeFamily\_Noctuidae, FoodPlant\_*, *MedicalPlant\_*, *MultiUsePlSp\_*,  
*MultiUsePlGe\_*, *MultiUsePlFa\_* ⇒ *MedicalCrop\_X (OT2)*

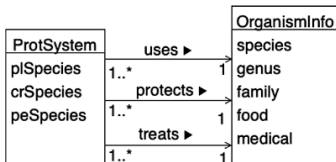
How to increase the readability of the implications?



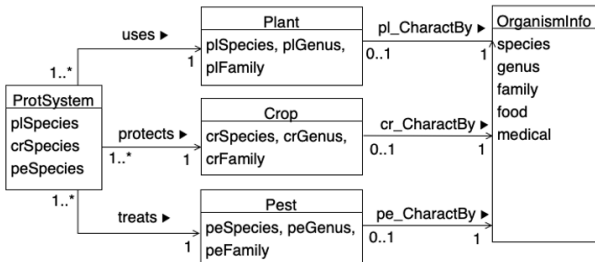
# Data structured in several tables

Split data: the TwoTables and Relational models

## b) TwoTables



## a) Relational

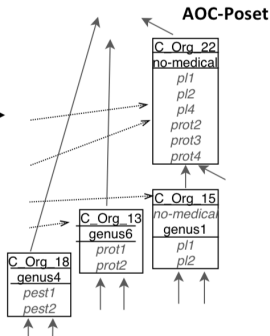


### Formal Concept Analysis

Formal Context

<i>OrganismInfo</i>	genus1	...	genus4	...	genus6	...	no-medical	...
plant1 (pl1)	x	...		...		...	x	...
plant2 (pl2)	x	...		...		...	x	...
prot1		...		...	x	...	x	...
prot2		...		...	x	...	x	...
pest1		...	x	...		...		...
pest2		...	x	...		...		...
...	...	...	...	...	...	...	...	...

Build conceptual structures

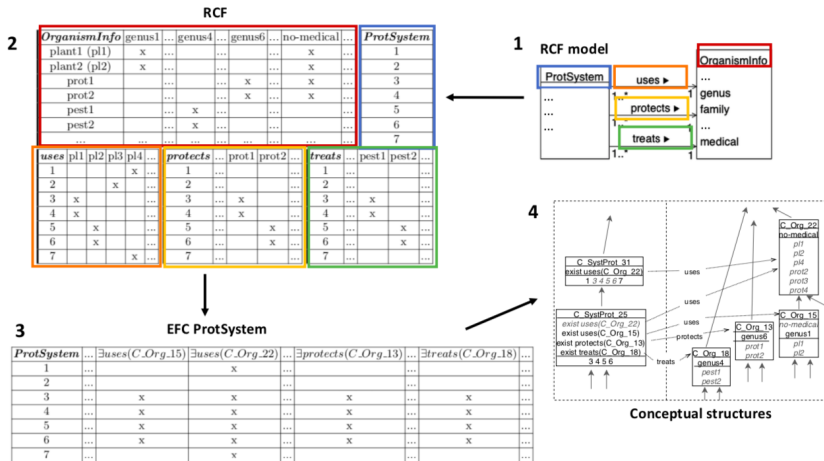


Ganter, B., Wille, R. Formal Concept Analysis. Mathematical Foundations. Springer. 1999

# Methodology based on Relational Concept Analysis (RCA)

RCA extends FCA

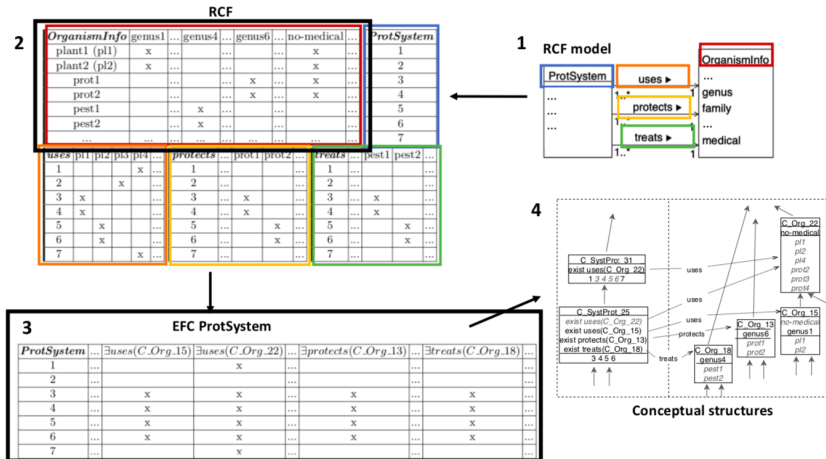
## Relational Concept Analysis



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

# Methodology based on Relational Concept Analysis (RCA)

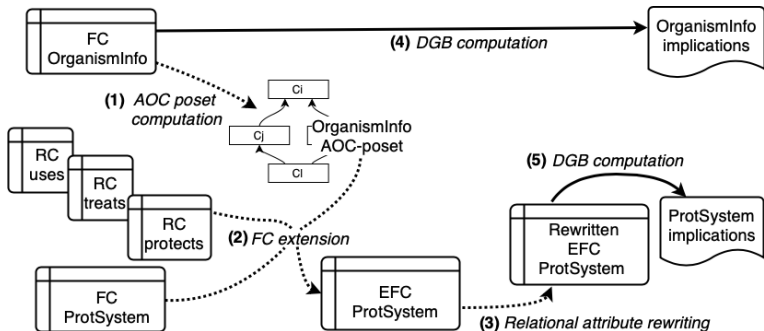
Example for the TwoTable model



We compute the DGB for FC OrganismInfo and EFC ProtSystem.



# Process for TwoTables model



The process can be generalized to any RCF model

### Quantitative description of the RCFs and conceptual structures (AOC-posets)

Representation	Formal context	#objects	#attributes	#relational attributes	#concepts AOC Poset (initial step)	#concepts AOC Poset (last step)
OneTable	FC CombinedRoles	721	1113	0	1005	1005
TwoTables	<i>All formal contexts</i>	1321	1078	2250	751	1750
	EFC ProtSystem	721	0	2250	1	1000
	FC OrganismInfo	600	1078	0	750	750
Relational	<i>All formal contexts</i>	1927	2169	3017	1507	2517
	EFC ProtSystem	721	0	767	1	1000
	EFC Plant	562	1008	750	700	705
	EFC Pest	29	45	750	36	37
	EFC Crop	15	38	750	20	25
	FC OrganismInfo	600	1078	0	750	750

## Implications (implic.) from the Duquenne-Guigues basis per scope (S) and maximum scope (Smax)

Representation	Formal context	#implic. S = 0	#implic. S = 1	#implic. S = 2	#implic. S = 3	#implic. S = [4-10] (avg)	#implic. S > 10 (avg)	#total implic. S > 0	Smax (#implic.)
OneTable	FC CombinedRoles	3360	1105	281	125	236 (33.71)	173 (2.34)	1920	721 (1)
TwoTables	EFC ProtSystem	4891	827	234	95	165 (23.57)	74 (1.90)	1395	721 (1)
	FC OrganismInfo	6069	1007	76	37	42 (7.00)	6 (1.2)	1168	35 (1)
	All FCs	10960	1834	310	132	207	80	2571	
Relational	EFC ProtSystem	3414	825	234	95	164 (23.42)	73 (1.87)	1391	721 (1)
	EFC Plant	5698	1509	132	64	85 (14.17)	25 (2.08)	1815	87 (2)
	EFC Pest	855	67	8	0	4 (2)	1	80	29 (1)
	EFC Crop	740	58	8	4	0	0	70	3 (4)
	FC OrganismInfo	6069	1007	76	37	42 (7)	6 (1.2)	1168	35 (1)
	All FCs	16776	3466	450	208	295	105	4532	



### OrganismInfo

$Family\_Meliaceae \implies NoFood$

The Meliaceae are not consumed

$Genus\_Salvia \implies Family\_Lamiaceae$

Genus Salvia implies Family Lamiaceae

### ProtSystem

$\exists protects(NoFood), \exists protects(Medical), \exists treats(Family\_Noctuidae)$   
 $\exists treats(NoFood), \exists treats(NoMedical) \implies \exists protects(Family\_Euphorbiaceae),$   
 $\exists protects(Species\_RicinusCommunis\&Genus\_Ricinus), \exists treats(Genus\_Spodoptera),$   
 $\exists treats(Species\_SpodopteraLitura), \exists uses(NoFood), \exists uses(NoMedical),$   
 $\exists uses(Family\_Asteraceae), \exists uses(Species\_WollastoniaDentata\&Genus\_Wollastonia)$   
**(TT1)**

$\exists treats(NoFood), \exists treats(NoMedical), \exists treats(Family\_Noctuidae),$   
 $\exists treats(Genus\_Spodoptera), \exists uses(NoFood), uses(NoMedical) \implies$   
 $\exists protects(Medical)$  **(TT2)**

# Implication rules

## Examples for the Relational model

### ProtSystem

$$\begin{aligned} & \exists \text{treats}(\text{PeFamily\_Noctuidae}), \exists \text{treats}(\text{PeGenus\_Spodoptera}), \\ & \exists \text{uses}(\text{pl\_CharactBy}(\text{NoFood})), \exists \text{uses}(\text{pl\_CharactBy}(\text{NoMedical})) \quad \implies \\ & \exists \text{protects}(\text{cr\_CharactBy}(\text{Medical})) \quad \textbf{(Rel2)} \end{aligned}$$

### Crop

$$\begin{aligned} & \exists \text{cr\_CharactBy}(\text{NoFood}), \exists \text{cr\_CharactBy}(\text{Medical}) \quad \implies \\ & \text{CrSpecies\_RicinusCommunis}, \\ & \text{CrGenus\_Ricinus}, \text{CrFamily\_Euphorbiaceae}, \exists \text{cr\_CharactBy}(\text{Family\_Euphorbiaceae}), \\ & \exists \text{cr\_CharactBy}(\text{Species\_RicinusCommunis\&Genus\_Ricinus}) \quad \textbf{(Rel1)} \end{aligned}$$

Crops, used in medical care and not consumed, are from the Ricinus Communis species

### Plant

$$\begin{aligned} & \text{PlFamily\_Poaceae}, \exists \text{pl\_CharactBy}(\text{NoMedical}), \exists \text{pl\_CharactBy}(\text{Family\_Poaceae}) \\ & \implies \exists \text{pl\_CharactBy}(\text{NoFood}) \end{aligned}$$

Family Poaceae plants not used in medical care, are also not consumed

# Implication rules

Example of implications for the three models

## OneTable

$PeFamily\_Noctuidae, MedicalCrop, NoFoodCrop \implies PeGenus\_Spodoptera, NoFoodPlant, NoMedicalPlant, NoMultiUsePISp, NoMultiUsePIGe, NoMultiUsePIFa, NoPISplsCrEw, NoPIGelsCrEw, PIFamily\_Asteraceae, PeSpecies\_SpodopteraLitura, CrSpIsPIEw, CrGelsPIEw, PISpecies\_WollastoniaDentata, PIGenus\_Wollastonia, CrSpecies\_RicinusCommunis, CrGenus\_Ricinus, CrFamily\_Euphorbiaceae$  (OT1)

## TwoTables

$\exists protects(NoFood), \exists protects(Medical), \exists treats(Family\_Noctuidae), \exists treats(NoFood), \exists treats(NoMedical) \implies \exists protects(Family\_Euphorbiaceae), \exists protects(Species\_RicinusCommunis\&Genus\_Ricinus), \exists treats(Genus\_Spodoptera), \exists treats(Species\_SpodopteraLitura), \exists uses(NoFood), \exists uses(NoMedical), \exists uses(Family\_Asteraceae), \exists uses(Species\_WollastoniaDentata\&Genus\_Wollastonia)$  (TT1)

## Relational

$\exists cr\_CharactBy(NoFood), \exists cr\_CharactBy(Medical) \implies CrSpecies\_RicinusCommunis, CrGenus\_Ricinus, CrFamily\_Euphorbiaceae, \exists cr\_CharactBy(Family\_Euphorbiaceae), \exists cr\_CharactBy(Species\_RicinusCommunis\&Genus\_Ricinus)$  (Rel1)

OT1 and TT1 combine (in blue) information about RicinusCommunis as a protected crop and other information on protection systems, Rel1 contains only information about the (protected) crop role of RicinusCommunis

Dividing data into separate FCs, which introduces RCs:

- ❖ Produces more implications
- ❖ The implications are divided into coherent subsets, i.e. one per FC, simplifying their analysis
- ❖ Reduces the length of the rules
- ❖ Combining less attributes makes the implications easier to read
- ❖ Eliminate the technical attributes introduced in OneTable to make the dataset coherent
- ❖ The more data are separated, the more implications are understandable

## Conclusion

- ❖ Knomana dataset
- ❖ Split data
- ❖ Effect on the implications

## Future work

- ❖ Clean the rules to limit redundancies (Saoud et al, 2021)
- ❖ Classify the rules depending their content (Saoud et al, 2021)
- ❖ Use Iceberg lattice to reduce the number of relational attributes

Saoud J., Gutierrez A., Huchard M., Marnotte P., Silvie P., Martin P., 2021. Explicit versus Tacit Knowledge in Duquenne-Guigues Basis of Implications – Preliminary Results. Analyzing Real Data with Formal Concept Analysis (RealDataFCA'2021), <https://icfca2021.sciencesconf.org/resource/page/id/12>

Lina Mahrach



Pierre Silvie



Priscilla Keip



Pascal Marnotte



Pierre Martin



Alain Gutierrez



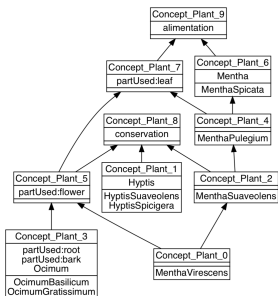
Marianne Huchard

# Thank you!

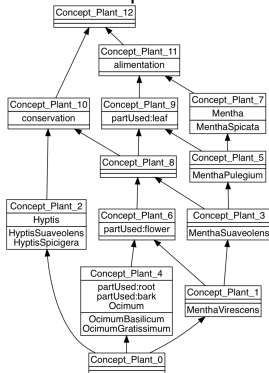
# Alternative conceptual structures

Plant	partUsed:flower	partUsed:leaf	partUsed:root	partUsed:bark	conservation	alimentation	Hyptis	Mentha	Ocimum
HyptisSuaveolens					x	x			
HyptisSpicigera					x		x		
MenthaPulegium		x				x		x	
MenthaSpicata						x			
MenthaSuaveolens		x			x	x		x	
MenthaVirescens	x	x			x	x		x	
OcimumBasilicum	x	x	x	x	x	x			x
OcimumGratissimum	x	x	x	x	x	x			x

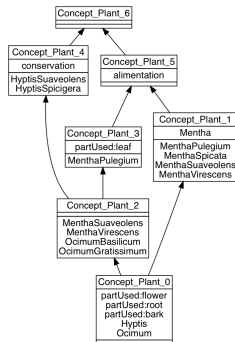
## AOC-poset

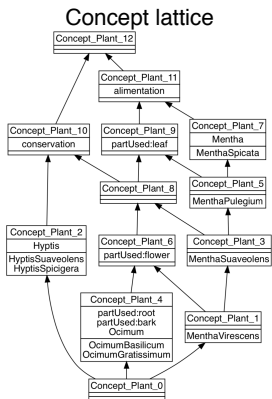


## Concept lattice



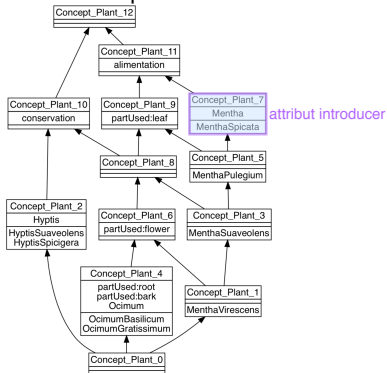
## Iceberg

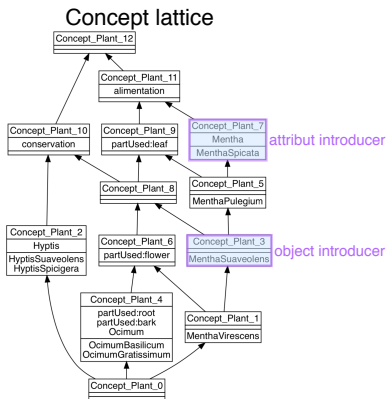






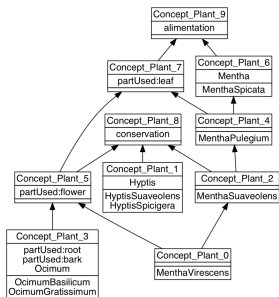
## Concept lattice





# Alternative conceptual structures

## AOC-poset



## Concept lattice

