Towards
(Re)construction of the Theory of
Linguistic Oppositions
within the Framework of Interactive Linguistics

André WLODARCZYK
wlodarczyk.andre@gmail.com

MAIN TOPICS

(1) Linguistic typology practice
(2) Interactive Linguistics
(3) Need for Interactive Modelling using the implementation of FCA in *Semana*-suite
(4) SEMANA tools
(5) One Example of Reconstruction
(6) Linguistic Signs
    (6.1) Binary oppositions
    (6.2) Many-valued oppositions
(7) Conclusions

Linguistic typology practice

"Form" and "Matter" in Structural Linguistics

	OBJECT	APPROACH	THEORY
FORM (Structures)	types universal homogeneous	Deductive Synthesis using rules	**L = (*W*, *G*)** Language is a set of sentences *generated* by grammar rules ***G*** from words ***W*** ***Prediction***
MATTER (Data)	instances specific heterogeneous	Inductive Analysis of analogies	**L = (*W*, *L*)** Language is a set of sentences ***L*** *analysed* as words ***W*** ***Explanation***

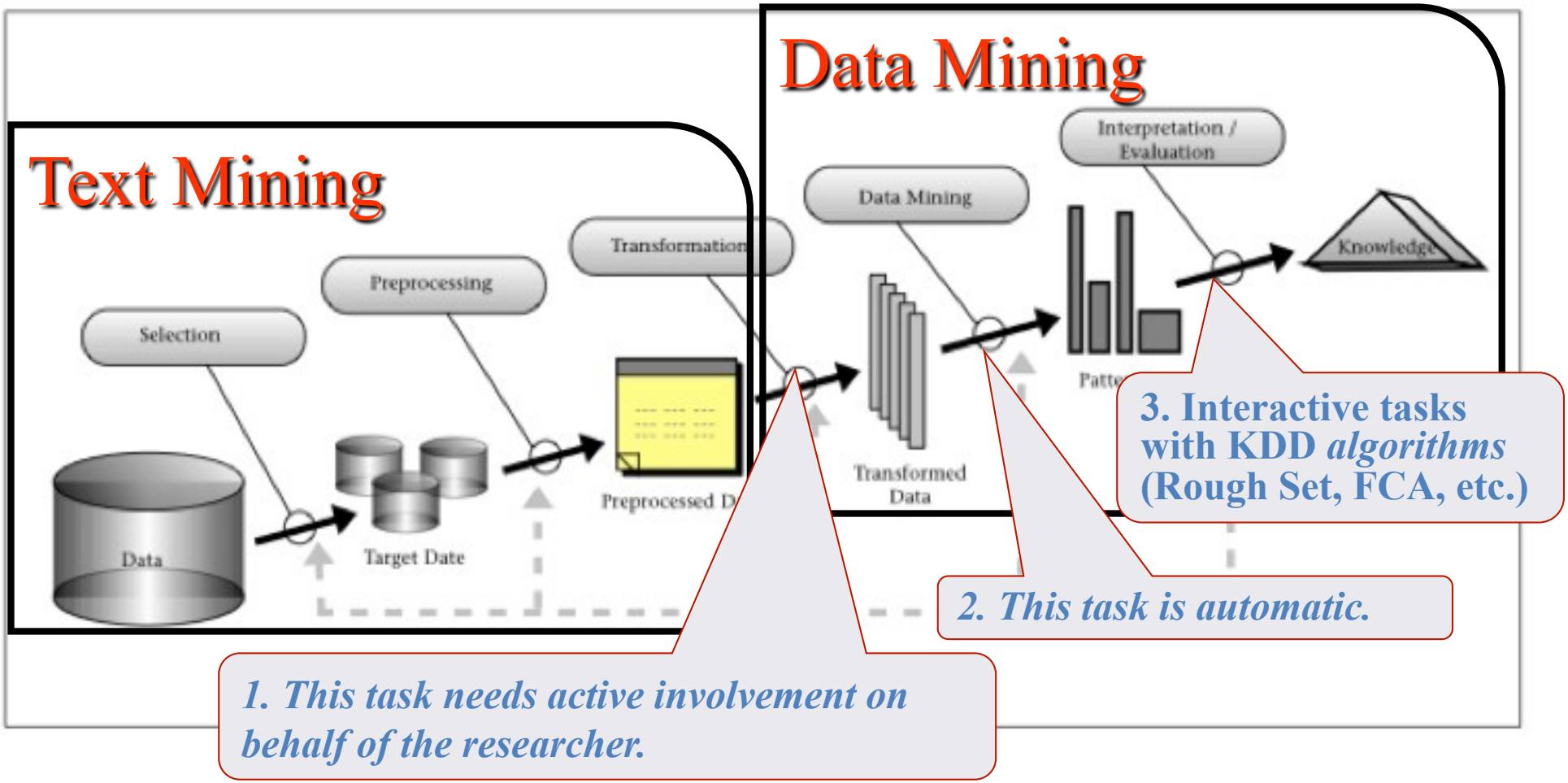
ALTMAN G. (1987) "The Levels of Linguistic Investigation", *Theoretical Linguistics*, vol. 14, edited by H. Schnelle, W. de Guyter, Berlin - New York

Structural and Computational Linguistics

	Structural Linguistics	**Computational Linguistics**
FORM (Structures)	**THEORY-oriented Linguistics** (Formal Generative Linguistics)	**Natural Language Processing** (Lexicon-Functional Grammars, Unification Grammars, Logic Grammars)
MATTER (Data)	**DATA-oriented Linguistics** (Linguistic Typology)	**Human Language Technology** (Corpus Linguistics, Lexicons and Thesauri - WordNet, FrameNet etc.)

INTERACTIVE LINGUISTICS

Text Mining and Data Mining



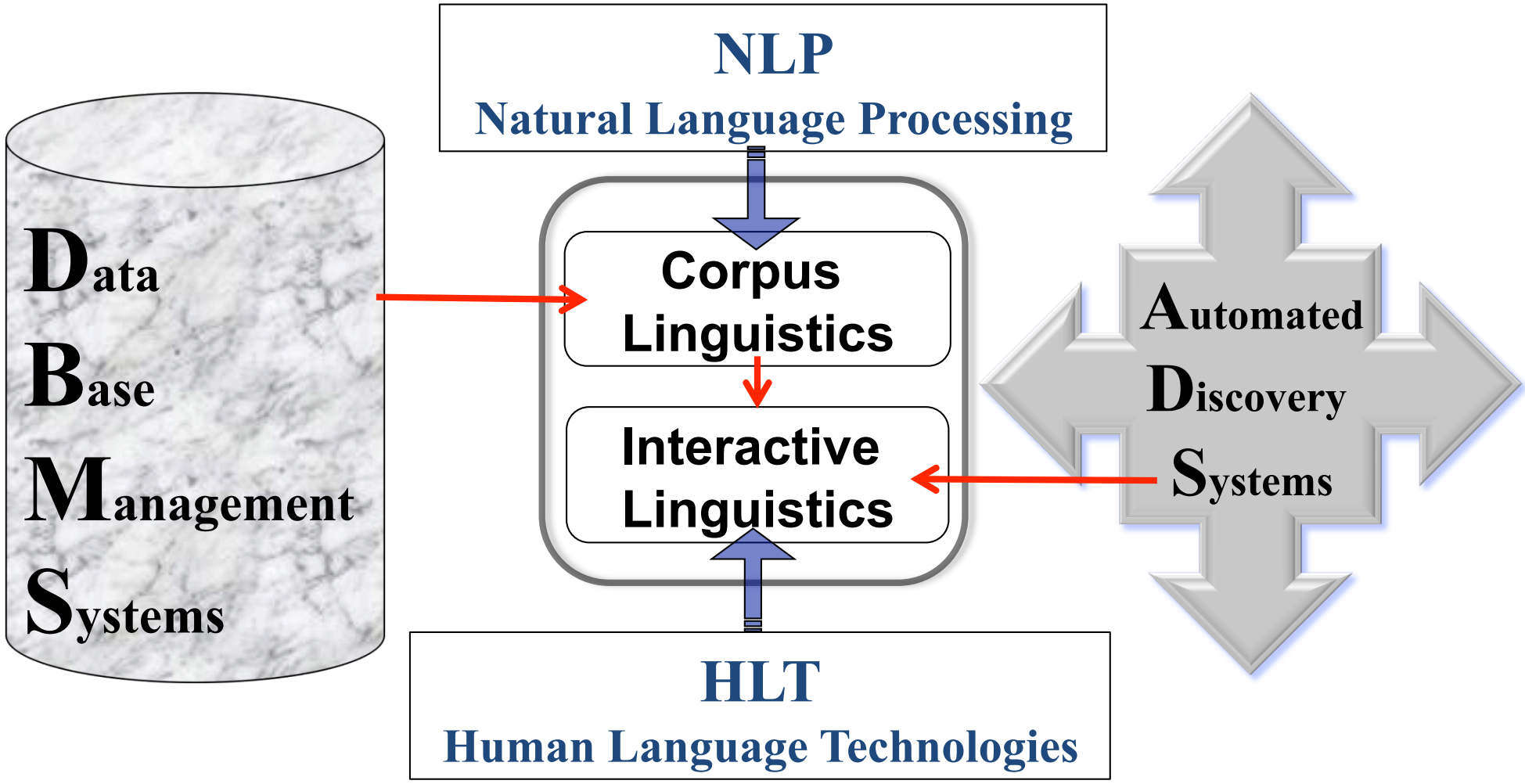
From Data Mining to Knowledge Discovery in Databases by Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, AI Magazine 1997 (American Association for Artificial Intelligence)

OBJECTS – APPROACHES - TASKS

Objects	**Text Data**	**Symbolic Data**
Approaches	**Corpus Linguistics** Text Document Exploration (Text Mining)	**Interactive Linguistics** Linguistic Knowledge Extraction (Data Mining)
Tasks	1. Selection 2a. Preprocessing 2b. Filtering	3. Transformation 4. Analysis 5. Evaluation

INTERACTIVE LINGUISTICS

In language studies Interactive Linguistics extends Text Mining using Symbolic Analysis (Data Mining) tools.



SEMANA-suite

SEMANA

The architecture of SEMANA was conceived by ***André WLODARCZYK*** and implemented in *Transcript* ® (an object-oriented programming language) for Windows, Apple and Linux platforms by ***Georges SAUVET*** and ***André WLODARCZYK***.

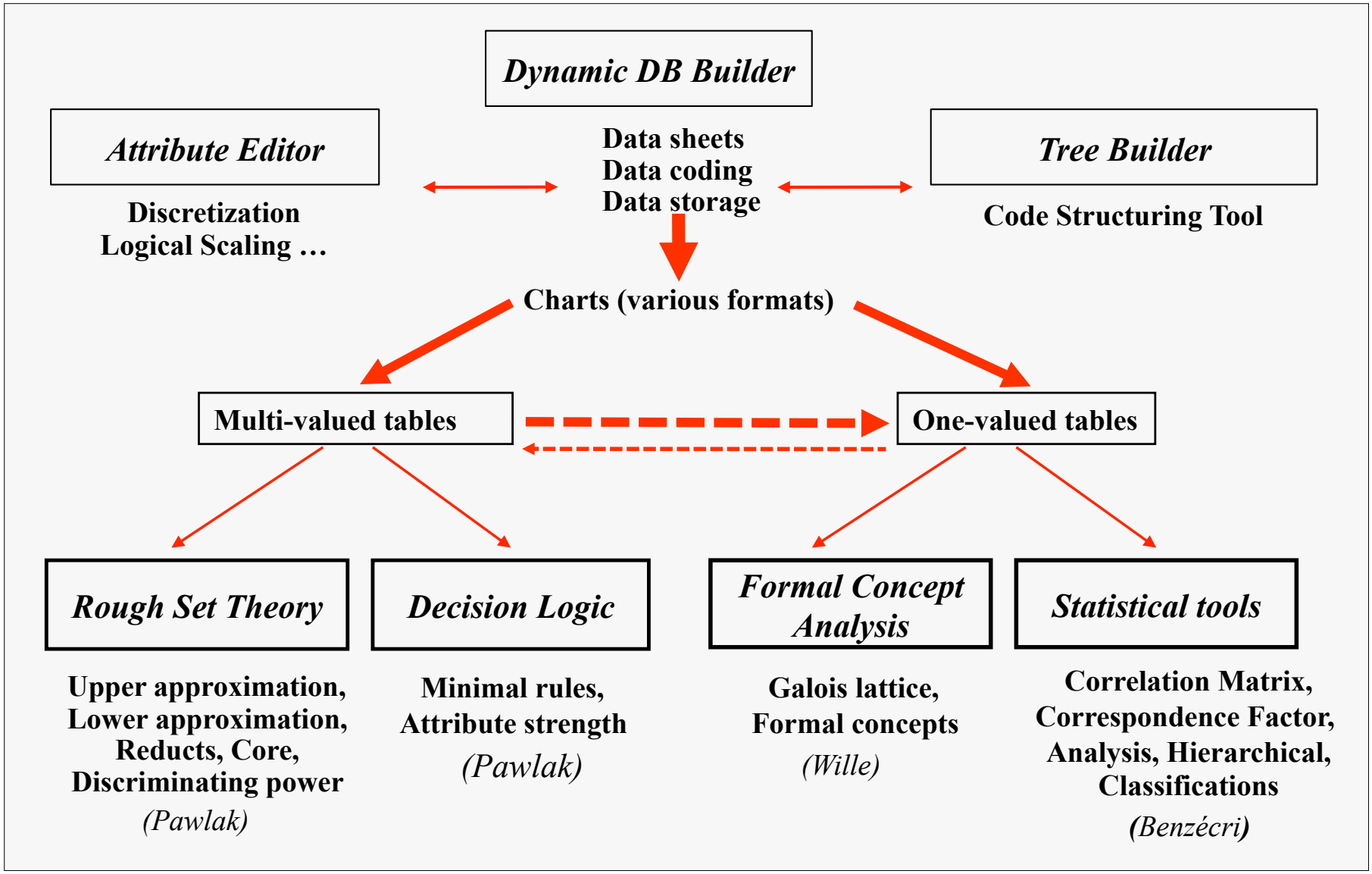
The symbolic processing tools are property of the authors of the following theories : FCA (Formal Concept Analysis), RSA (Rough Set Analysis) and DL (Decision Logic). They were implemented by ***Georges SAUVET***.

Statistical tools (**STAT 3**) were implemented by ***Georges SAUVET*** using Benzécri's algorithms (*originally* written in *Fortran*).

Some algorithms (such as the calculators of Core Concept, Central and Master Concept, Intensional and Extensional "semions") are property of **Georges Sauvet** and **André Wlodarczyk**.

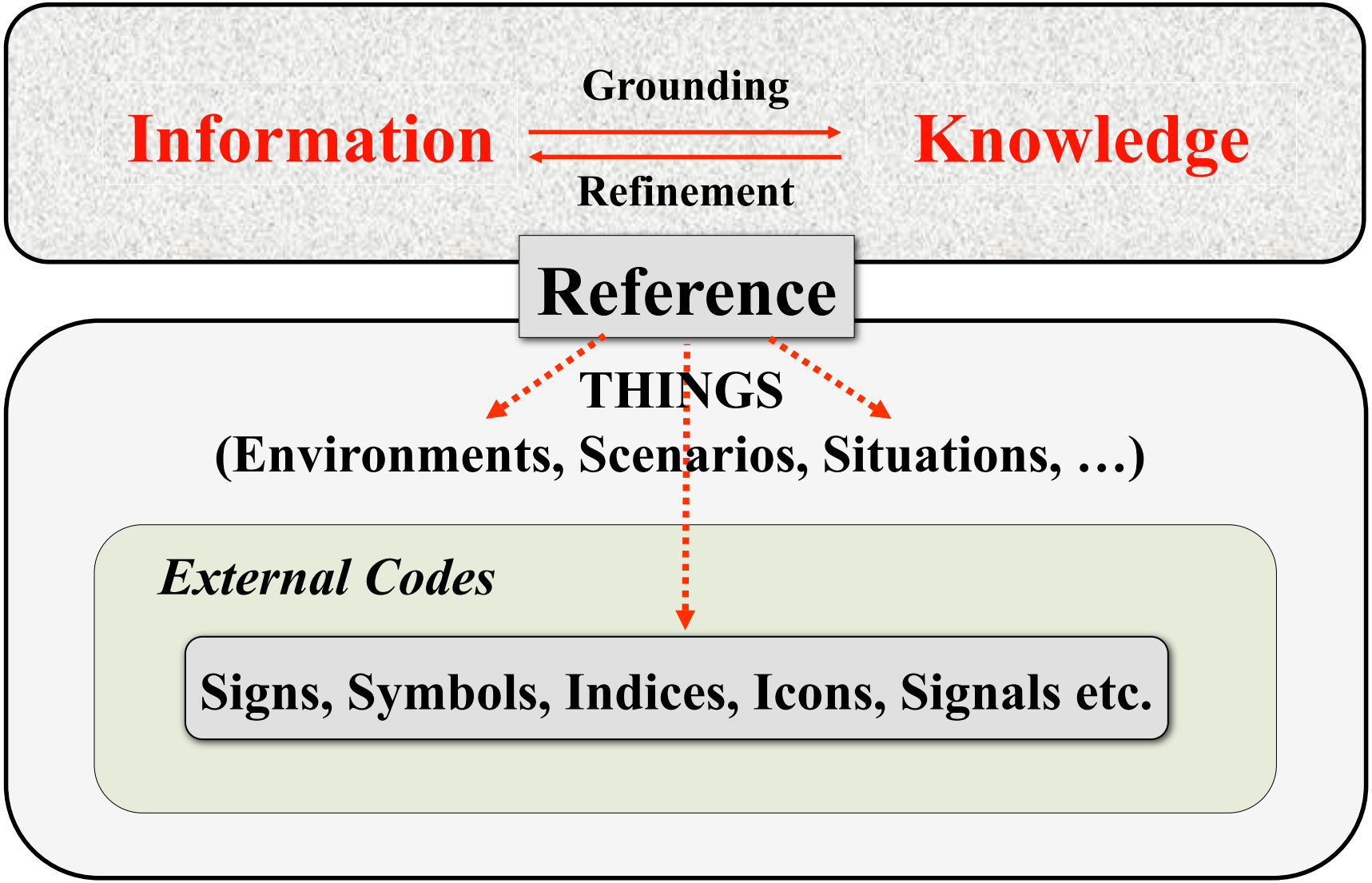
Architecture of *SEMANA*

Software-suite for Apple, Windows and Linux



Linguistic Signs

SEMIOSIS


Information
Grounding
Refinement
Knowledge
Reference
THINGS
(Environments, Scenarios, Situations, …)
External Codes
Signs, Symbols, Indices, Icons, Signals etc.

Sign Usage and Sense

The **Sign** is a **structure** with usages U as objects and descriptions F as formulae.

$$\textbf{\textit{Sign}} = < U,F >$$

Let F be a set of atomic formulae $F=\{\phi, \chi, \psi...\}$ and let Φ be a subset of formulae in F (i.e.: $\Phi \subseteq F$). Let $X=\{x, y, z...\}$ be a subset of U (i.e.: $X \subseteq U$).

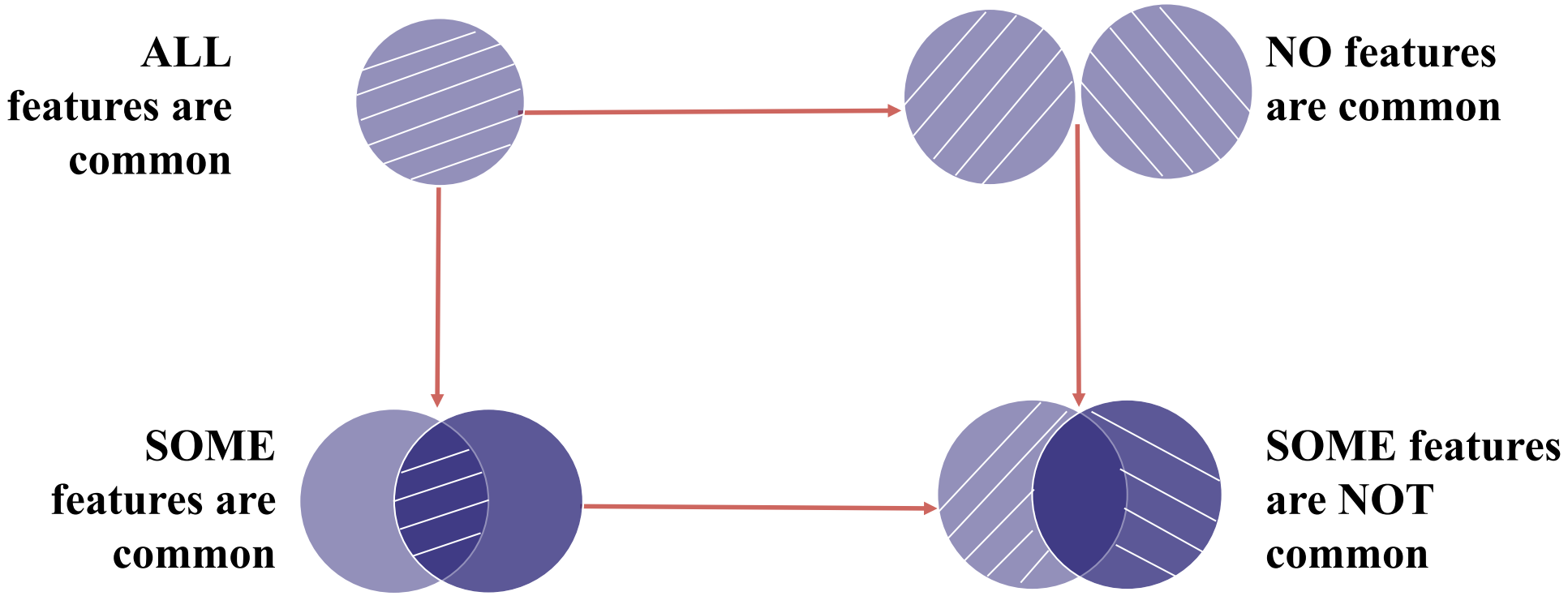
The ***Usage*** of a sign is defined as its extension; i.e.: a typified set (class) of uses.

$$\|\Phi\|_{\text{Sign}} = \{x \in X : x \vDash_{Sign} \Phi\}$$

The ***Sense*** of a sign is defined as its intension.

$$\|X\|_{\text{Sign}} = \{\phi \in \Phi : \phi \vDash_{Sign} X\}$$

Similarity & Distinction

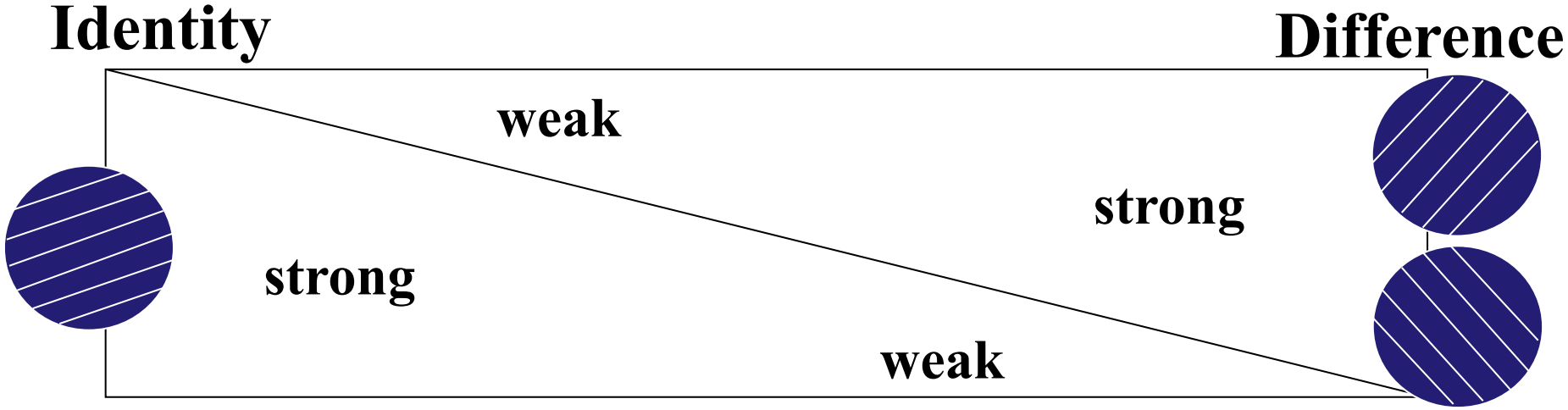


Following some definitions of 'Similarity' and 'Distinction' by Jerzy Pogonowski (1991), Linguistic Oppositions, *UAM Scientific Editions, Poznań, pp. 125*

SIMILARITY AND DISTINCTION

Linguistic signs can be compared within dual *continuous* spaces which have *identity* and *difference* as their extreme cases .

Morphemes oppose in pairs of similarity and distinction.



	Similarity	**Distinction**
Close Senses	strong	weak
Distant Senses	weak	strong

One Example of Reconstruction

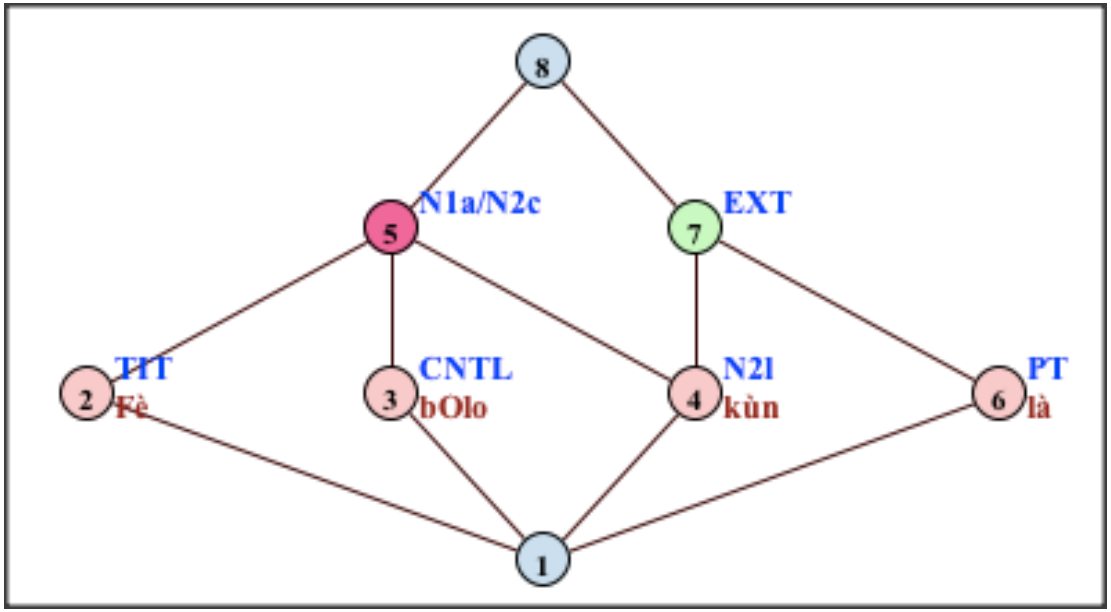
Possession in Bambara (data 1)

	lá	autres		
PRE_1	SN_1 *partie de* SN_2	SN_1 ≠ *partie de* SN_2		
PRE_2		ANIME		
PRE_3		CONCRET		
		kùn	autres	
PRE_4		SN_2 LOCATIF de SN_1		
			bólo	fè
ASS_1	EXISTENCE	EXISTENCE		
ASS_2			CONTROLE	
ASS_3				TITRE

NOTE SUR L'EXPRESSION ENONCIATIVE DE LA POSSESSION EN BAMBARA, Haimund Kastenholz

Possession in Bambara (data 2)

FCA	CNTL	EXT	N1a	N2c	N2l	PT	TIT
là	*	x	*	*	*	x	*
kùn	*	x	x	x	x	*	*
bÒlo	x	*	x	x	*	*	*
Fè	*	*	x	x	*	*	x

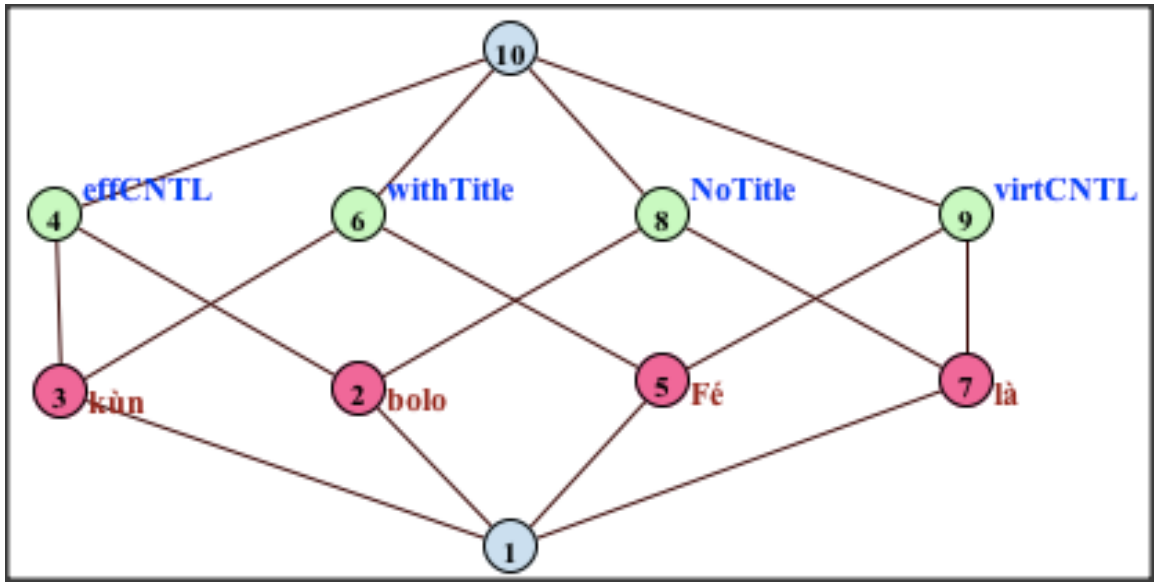


NOTE SUR L'EXPRESSION ENONCIATIVE DE LA POSSESSION EN BAMBARA, Haimund Kastenholz

Possession in Bambara (proposal)

fca	effCNT	virtCNT	withTitl	NoTitle
là	0	x	0	x
Fé	0	x	x	0
kùn	x	0	x	0
bÒlo	x	0	0	x
total	2	2	2	2

Virtual Control without Title
Virtual Control with Title
Effective Control with Title
Effective Control without Title



NOTE SUR L'EXPRESSION ENONCIATIVE DE LA POSSESSION EN BAMBARA, by Haimund Kastenholz
Reconstruction by André Wlodarczyk

Linguistic Oppositions

Structural linguists used to distinguish between 3 kinds of feature oppositions:

***privative* (binary), *equipollent* (multi-valued) and *gradual* (degree-valued)**

The value of the privative opposition is known as 'marked' in at least 3 ways:

1. as a 'positive' feature (present attribute) vs. a 'negative' feature (absent attribute), ex. Past tense (w.r.t. Present tense which is 'unmarkded')
2. as a 'distinguished" feature in one morpheme vs. two features within a unique morpheme (one of them being inverse to the former and the other one being the feature of their hypernym, ex. *woman* in the hierarchy

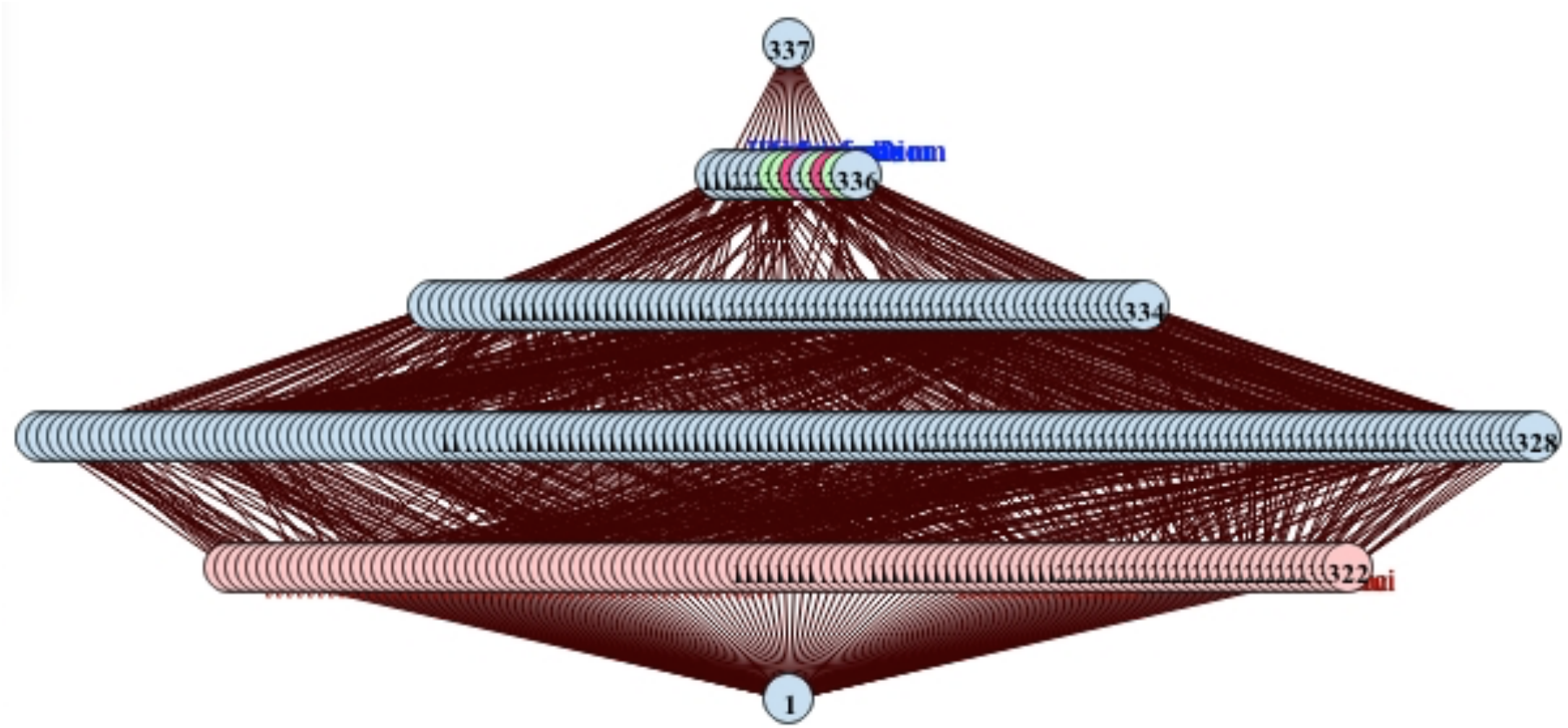
   (man, (man, woman))

1. as a 'neutralised' or 'irrelevant' feature in a morpheme which even though belonging to the given grammatical paradigm does not exhibit the expected feature, ex. in the pronoun 'I' where GENDER (mas and fem) is irrelevant.

MULTI-LAYERED LATTICE DIAGRAM

Clusters of 14 Polish morphemes described by 4 features:

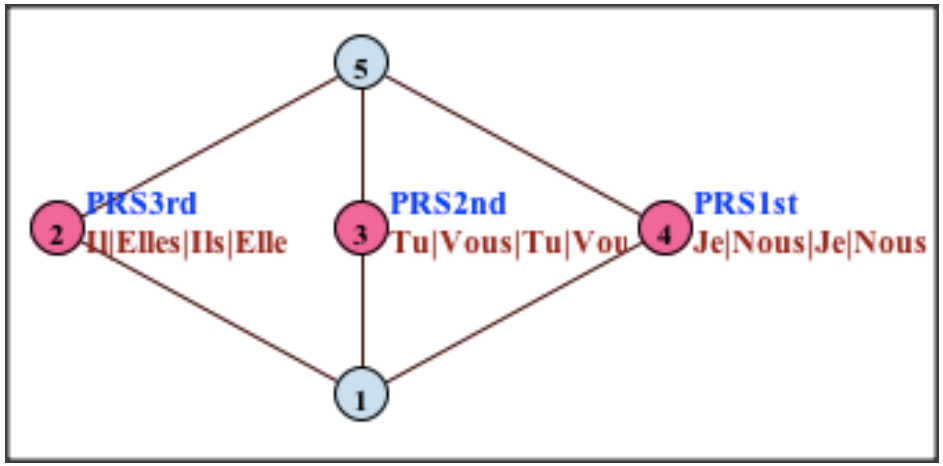
Case = {Nom, Gen, Dat, Acc, Ins, Loc}
Gender = {mas, fem, neu}
Number = {sin, plu}
Animacy = {ani, ina ,hum}

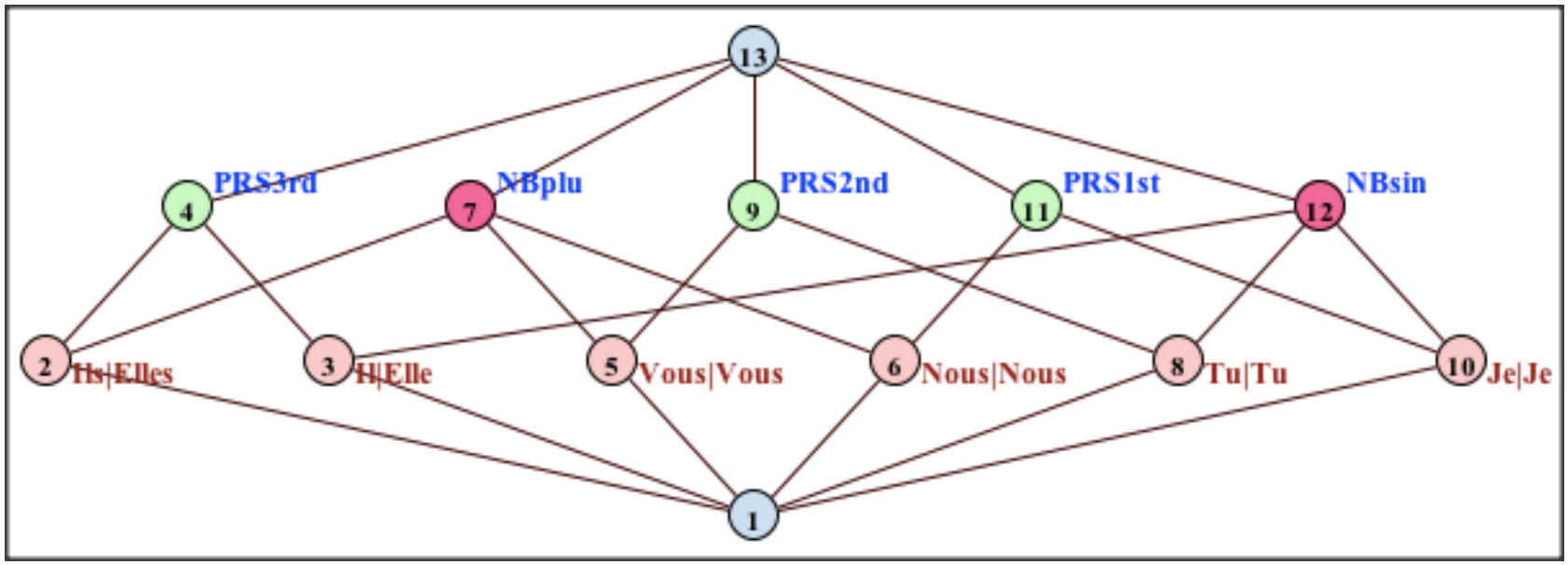


PERSONAL PRONOUNS (DATA)

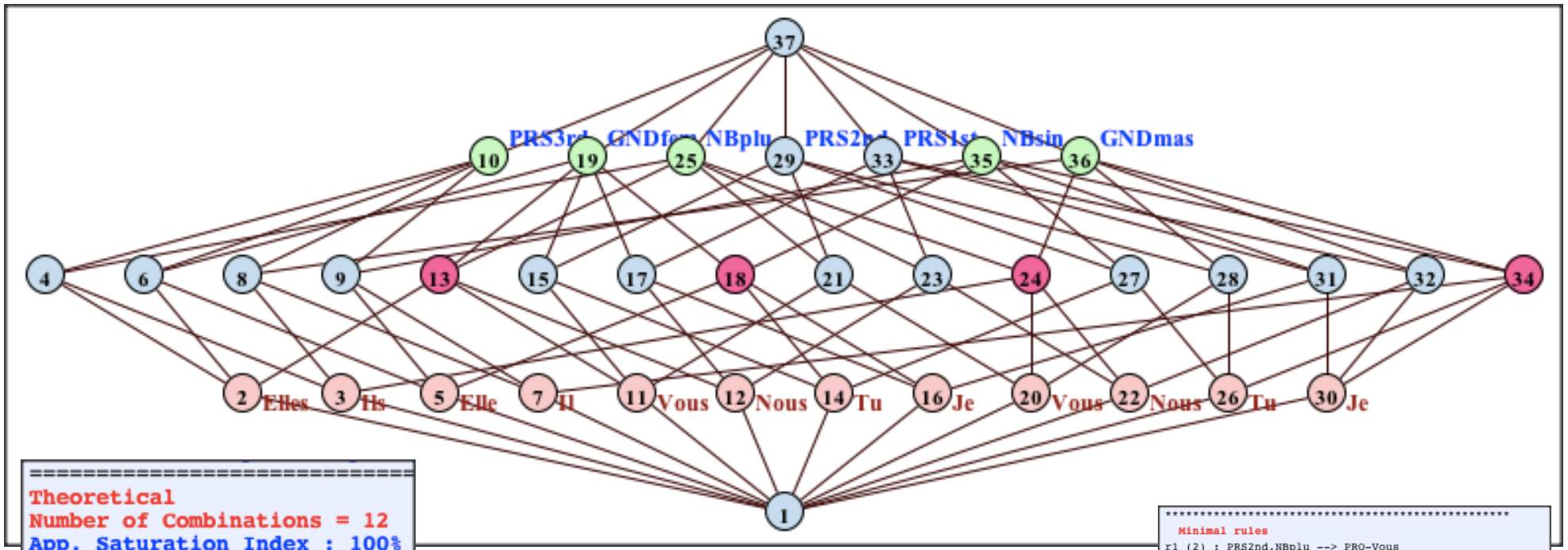
AV	PRS	NB	GND
Je	1st	sin	mas
Tu	2nd	sin	mas
Nous	1st	plu	mas
Vous	2nd	plu	mas
Je	1st	sin	fem
Tu	2nd	sin	fem
Nous	1st	plu	fem
Vous	2nd	plu	fem
Il	3rd	sin	mas
Elle	3rd	sin	fem
Ils	3rd	plu	mas
Elles	3rd	plu	fem

Personal Pronouns (1 & 2 Attributes)





Personal Pronouns (3 Attributes)




```
==============================
Theoretical
Number of Combinations = 12
App. Saturation Index : 100%
==============================
STATISTICAL USE OF AV
Attr        Value        occur
GND         fem          6
GND         mas          6

NB          plu          6
NB          sin          6

PRS         1st          4
PRS         2nd          4
PRS         3rd          4
==============================
```

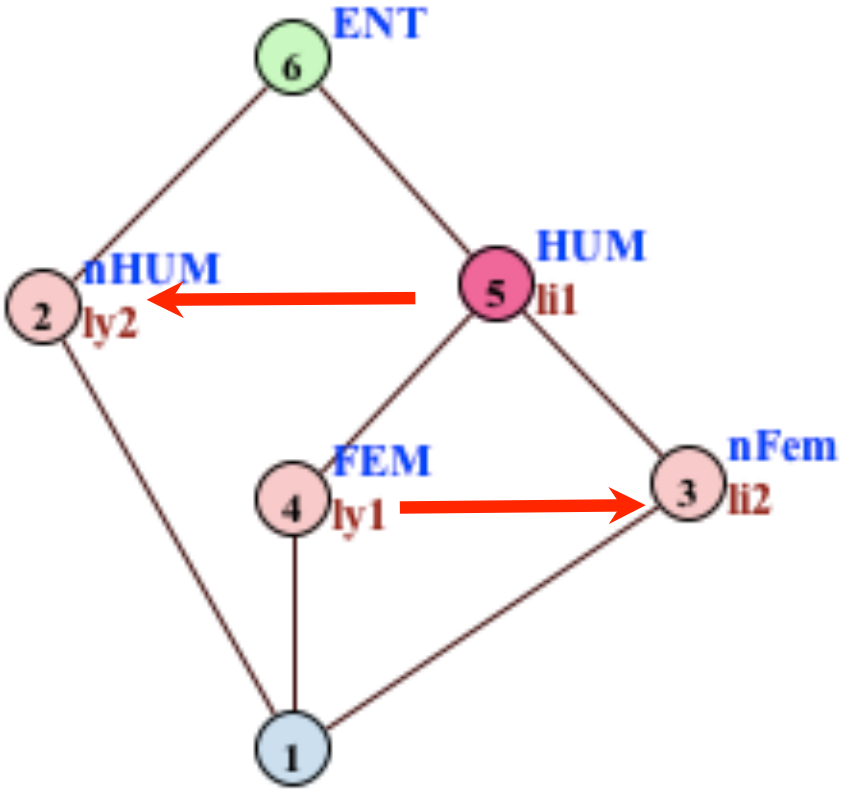
```
**************************************
CENTRAL FORMAL CONCEPT :
C13   {Nous,Vous,Elles},{NBplu,GNDfem}
C18   {Je,Tu,Elle},{NBsin,GNDfem}
C24   {Nous,Vous,Ils},{NBplu,GNDmas}
C34   {Je,Tu,Il},{NBsin,GNDmas}
```

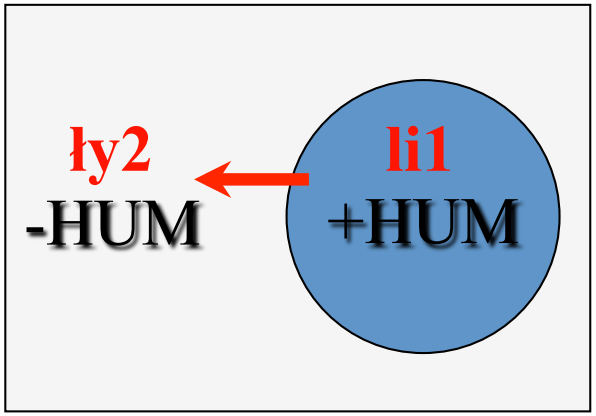
```
  Support of Rules
r1 : 2/PRO-Vous {Vous.Vous}
r2 : 2/PRO-Tu {Tu.Tu}
r3 : 2/PRO-Nous {Nous.Nous}
r4 : 2/PRO-Je {Je.Je}
r5 : 1/PRO-Il {Il}
r6 : 1/PRO-Ils {Ils}
r7 : 1/PRO-Elle {Elle}
r8 : 1/PRO-Elles {Elles}
```

```
**************************************************
  Minimal rules
r1 (2) : PRS2nd,NBplu --> PRO-Vous
r2 (2) : PRS2nd,NBsin --> PRO-Tu
r3 (2) : PRS1st,NBplu --> PRO-Nous
r4 (2) : PRS1st,NBsin --> PRO-Je
r5 (1) : PRS3rd,NBsin,GNDmas --> PRO-Il
r6 (1) : PRS3rd,NBplu,GNDmas --> PRO-Ils
r7 (1) : PRS3rd,NBsin,GNDfem --> PRO-Elle
r8 (1) : PRS3rd,NBplu,GNDfem --> PRO-Elles
**************************************************
  Support of Rules
r1 : 2/PRO-Vous {Vous.Vous}
r2 : 2/PRO-Tu {Tu.Tu}
r3 : 2/PRO-Nous {Nous.Nous}
r4 : 2/PRO-Je {Je.Je}
r5 : 1/PRO-Il {Il}
r6 : 1/PRO-Ils {Ils}
r7 : 1/PRO-Elle {Elle}
r8 : 1/PRO-Elles {Elles}
==========================================================
     Objects in the Table: 12
     Inconsistent Objects: 0
     Consistent Objects: 12
     Objects generated by the rules: 20
some objects are generated by several rules
**********************************************************
   Relative weight of attributes
              N      weight(%)
1.PRS         12      42.9
2.NB          12      42.9
3.GND          4      14.3
**********************************************************
```

Double Binary (“**Boomerang**”) Opposition

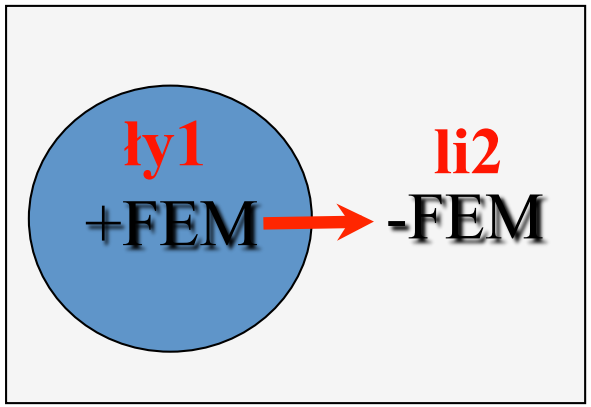
for Polish –li and –ły male/female past tense verb endings





Psy stały.
Pociągi stały.
Dzieci stały.

Ludzie stali.
Matka i dziecko stali.



Panie stały.

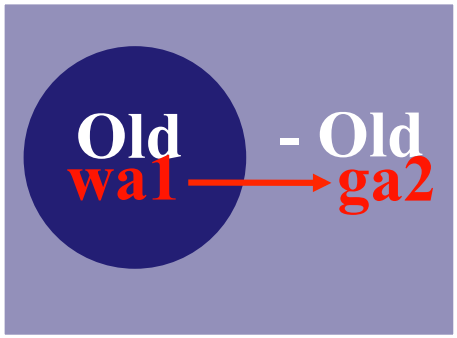
Panowie stali.

Inverse Opposition
between the Japanese ‘wa’ and ‘ga’ particles

In a BASE UTTERANCE :

-“GA” (ga1) is a marker of the Attention-driven Phrase (Subject with the status: ‘New’)

-“WA” (wa2) is a marker of the Attention-driven Phrase (Subject with status ‘not-New’)



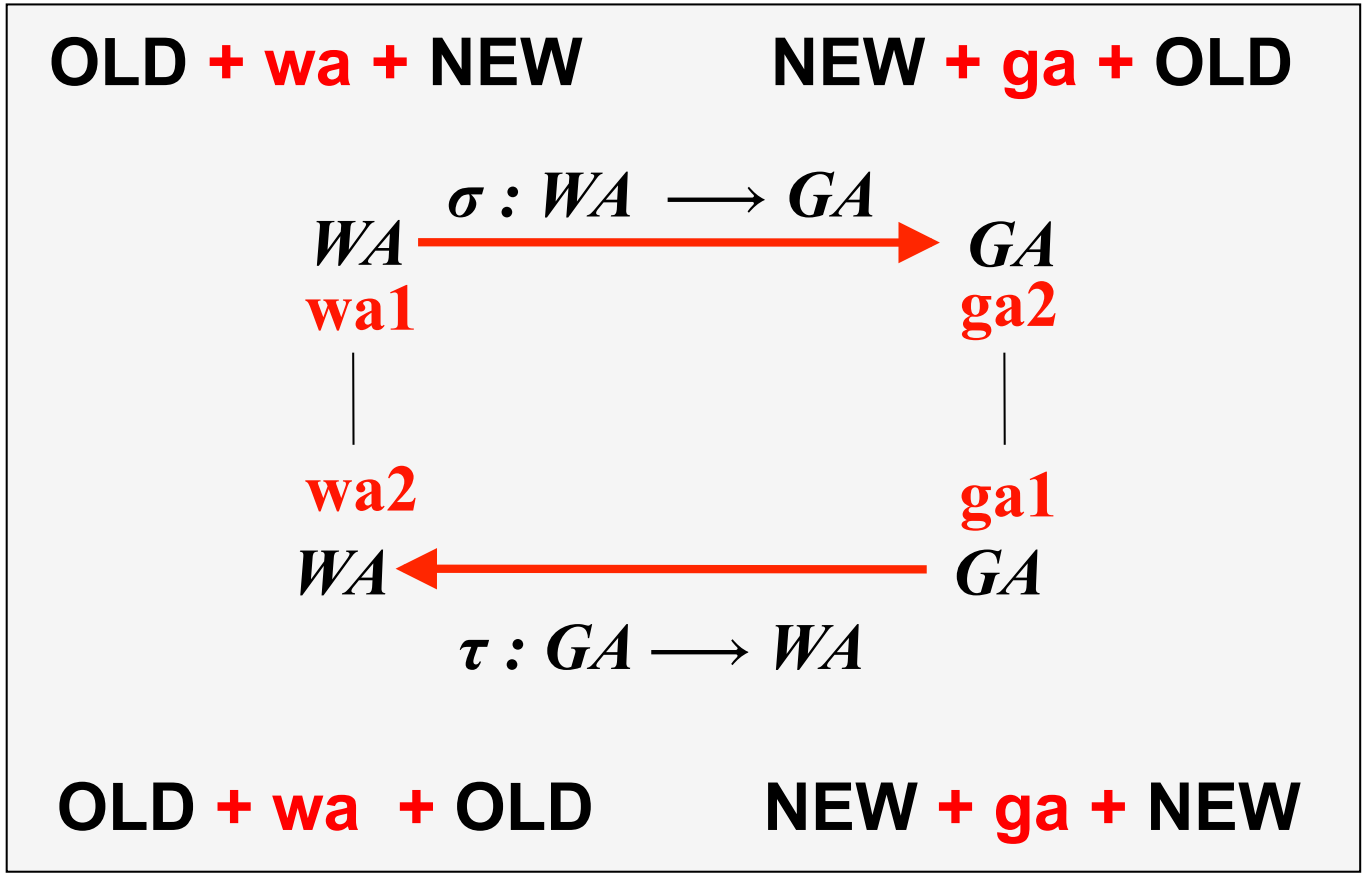
In an EXTENDED UTTERANCE :

-“WA” (wa1) is a marker of the Attention-driven Phrase (Topic with the status: ‘Old’)

-“GA” (ga2) is a marker of the Attention-driven Phrase (Focus with the status: ‘not-Old’)

Infomorphic Interpretation

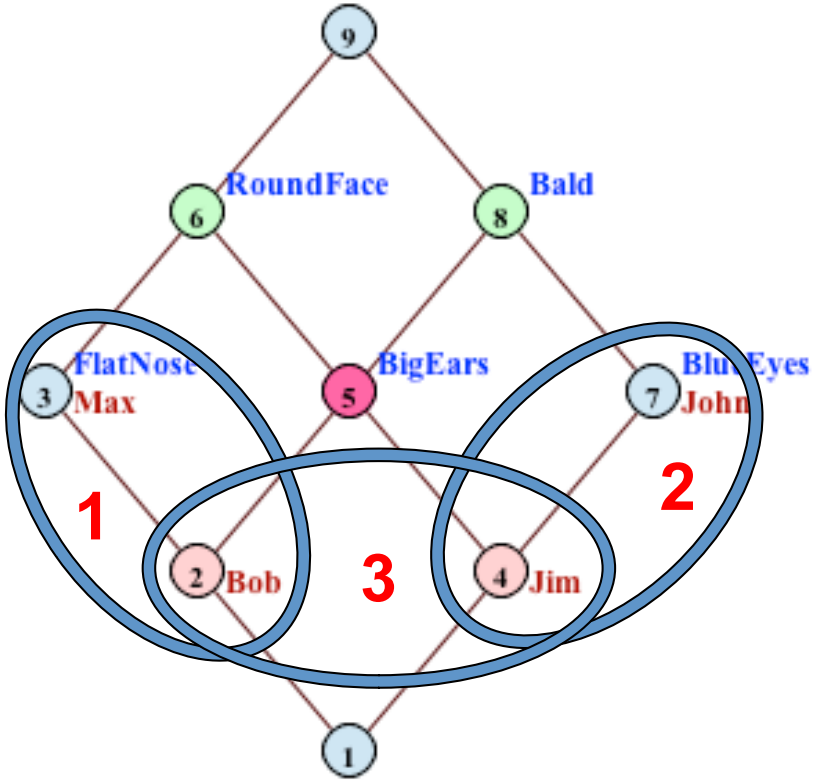
of the Opposition between the Japanese ‘wa’ and ‘ga’ particles



“Family Resemblance”

Multi-base Classes

fca	BigEars	BlueEyes	FlatNose	RoundFace	Bald
Jim	x	x	0	x	x
John	0	x	0	0	x
Bob	x	0	x	x	x
Max	0	0	x	x	0



CLASS STRUCTURE

Classe 1: Bob

C2 {Bob},{BigEars,FlatNose,RoundFace,Bald}

C3 {Bob,Max},{FlatNose,RoundFace}

Classe 2: Jim

C4 {Jim},{BigEars,BlueEyes,RoundFace,Bald}

C7 {Jim,John},{BlueEyes,Bald}

Classe 3: Jim,Bob

C5 {Jim,Bob},{BigEars,RoundFace,Bald}

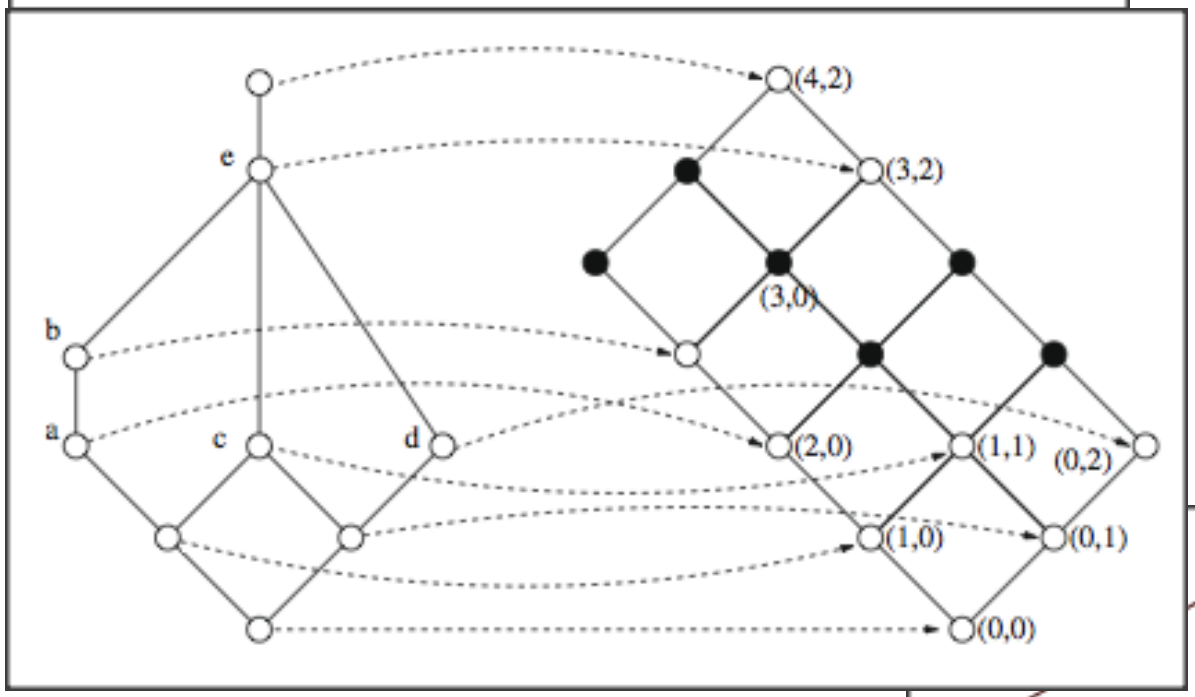
C6 {Jim,Bob,Max},{RoundFace}

C8 {Jim,John,Bob},{Bald}

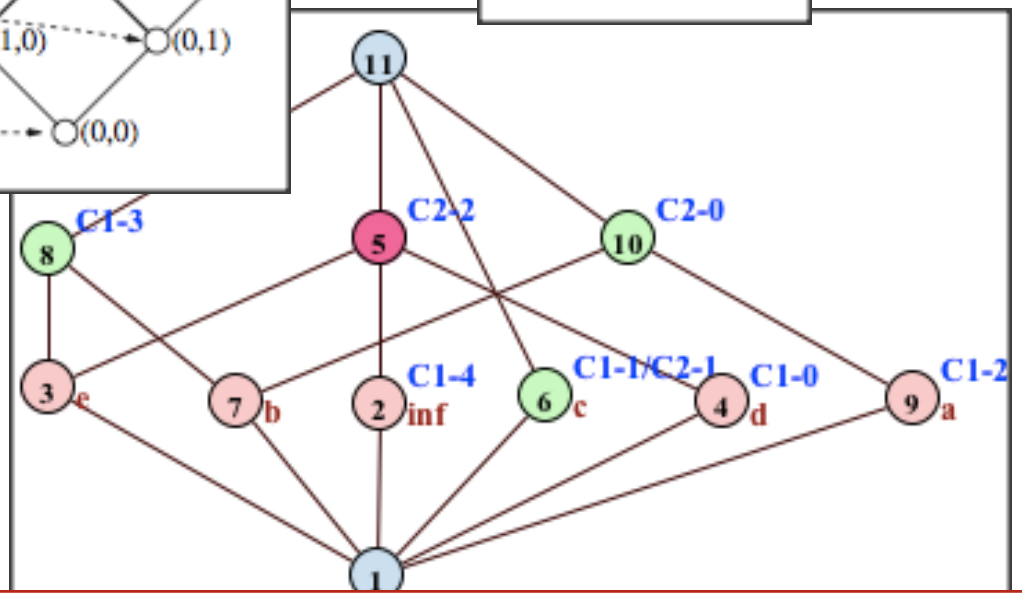
All Formal Concepts included

Ordinal and Nominal Many-Valued Attributes

Fig. 4. (a) A lattice L and a $\wedge$-homomorphism φ, (b) the many-valued relation R_L.

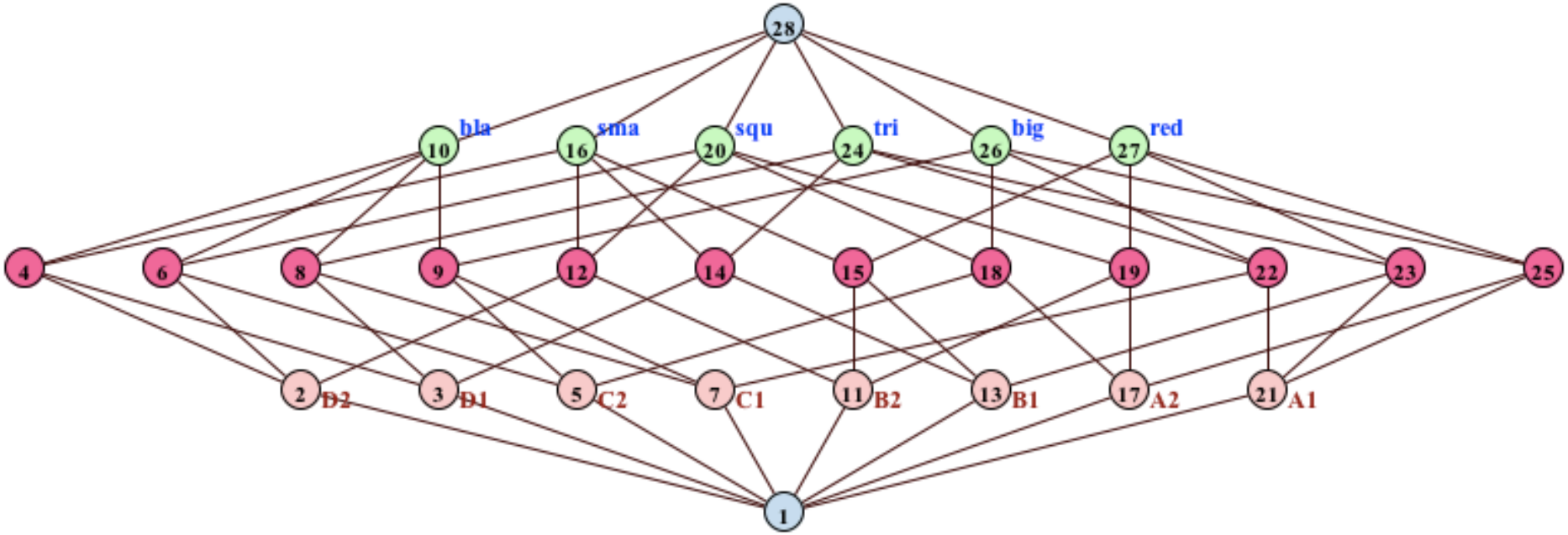


	C_1	C_2
a	2	0
b	3	0
c	1	1
d	0	2
e	3	2
T	4	2

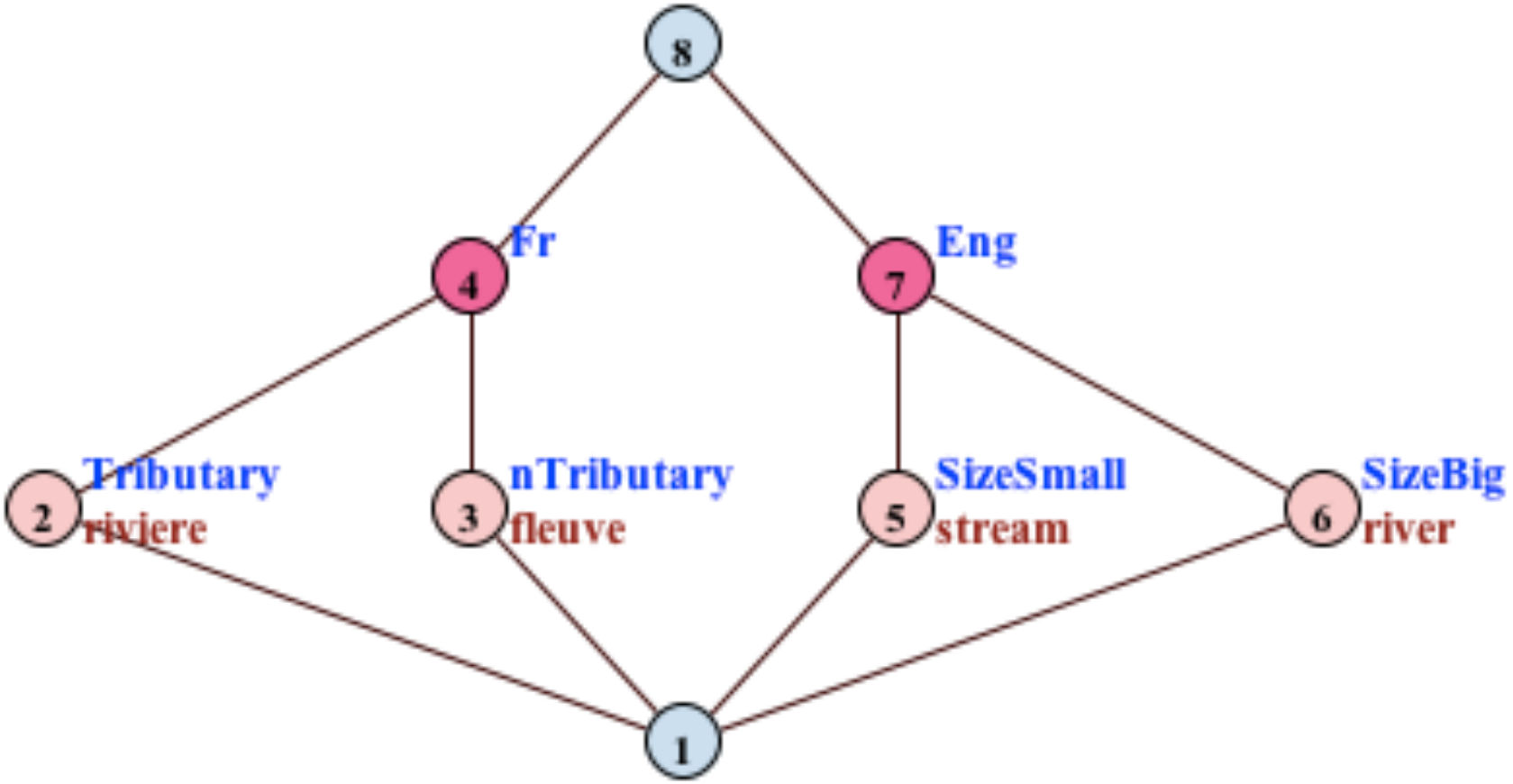


Representing lattices using many-valued relations by Alain Gély, Raoul Medina and Lhouari Nourine, published by Elsevier in "Information Sciences" 179 (2009) 2729–2739

Attributive Knowledge is similar to the Connectionist one



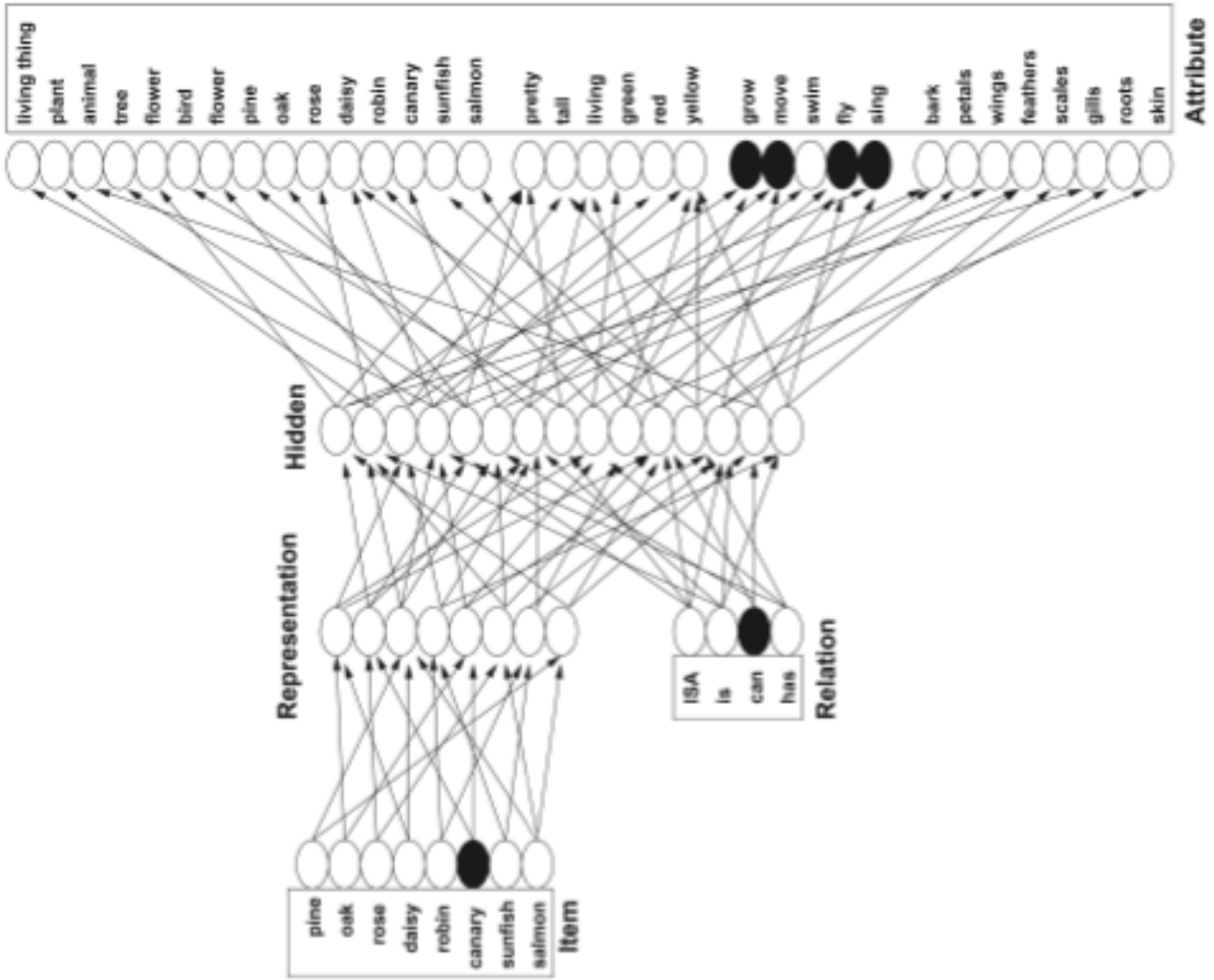
Attributive Knowledge



‘Ohio’ is big then it is said to be a “river” in English.

‘Ohio’ is tributary then it is said to be a “rivière” in French.

Connectionist (neural) Network

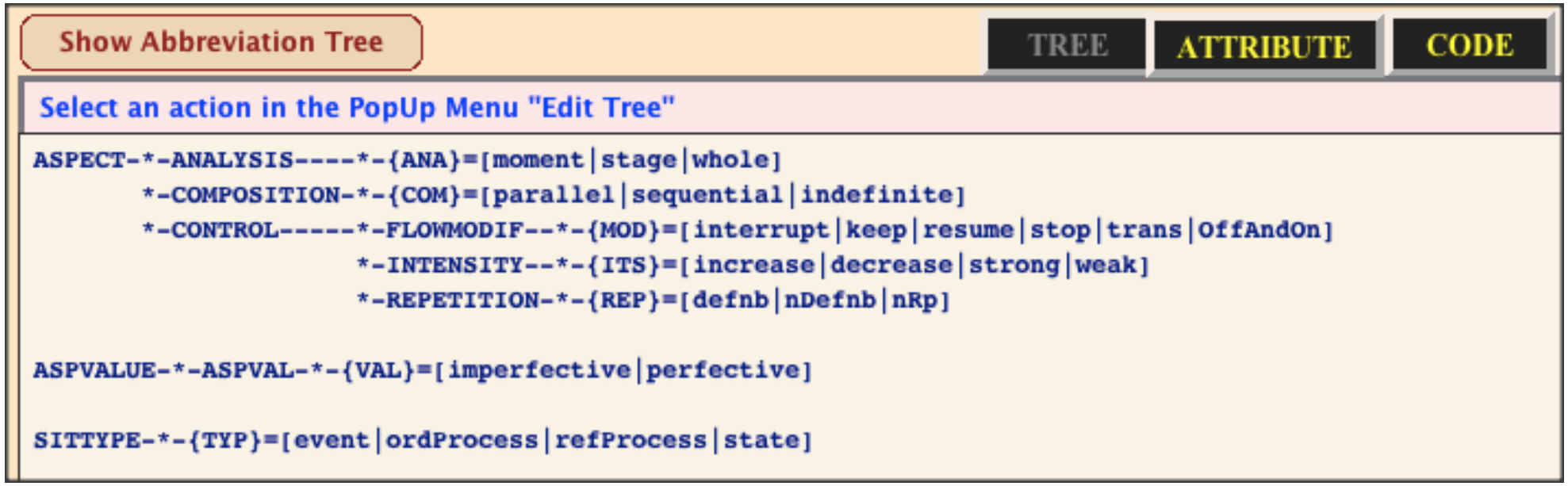


Rumelhart & Todd (1993)

Example of research on Polish Aspect (excerpt of data)

	AV	ANA	ANP	COM	ITS	MCP	MOD	REP	TYP	VAL
1	+AUX*asp	stage	before	*	*	prfImp	stop	nRp	ordProcess	imperfective
2	+ZERO	stage	before	*	*	*	stop	nRp	event	perfective
3	+SUF	stage	begin	*	*	prfImp	stop	nRp	ordProcess	imperfective
4	z	stage	after	*	*	impPrf	stop	nRp	ordProcess	perfective
5	za	moment	enter	*	*	impPrf	*	nRp	ordProcess	perfective
6	+AUX*asp	stage	begin	*	*	*	keep	nRp	ordProcess	perfective
7	+ZERO	stage	run	*	*	*	keep	nRp	ordProcess	imperfective
8	od	moment	finish	*	*	impPrf	trans	nRp	ordProcess	perfective
9	na	moment	finish	*	*	impPrf	*	nRp	ordProcess	perfective
10	na	stage	after	sequential	*	impPrf	*	nRp	ordProcess	perfective
11	+ZERO	stage	run	*	*	*	*	nDefnb	ordProcess	imperfective
12	prze	moment	finish	*	*	impPrf	*	defnb	ordProcess	perfective
13	prze	stage	after	parallel	*	impPrf	*	defnb	ordProcess	perfective
14	+ANL	stage	run	*	*	*	resume	nRp	ordProcess	perfective
15	do	moment	finish	*	*	impPrf	resume	nRp	ordProcess	perfective
16	+ZERO	stage	run	*	*	*	keep	nRp	ordProcess	imperfective
17	+AUX*asp	stage	run	*	*	*	interrupt	nRp	ordProcess	perfective
18	po	stage	run	*	*	impPrf	stop	nRp	ordProcess	perfective
19	na	moment	finish	*	*	impPrf	*	nRp	ordProcess	perfective
20	na	stage	after	parallel	*	*	*	nRp	ordProcess	perfective
21	prze	moment	finish	*	*	impPrf	trans	nRp	ordProcess	perfective
22	wy	moment	finish	*	strong	impPrf	*	nRp	ordProcess	perfective
23	+ZERO	moment	finish	*	strong	*	*	nRp	refProcess	perfective
24	+ZERO	stage	run	*	*	*	*	nRp	refProcess	imperfective
25	po+SUF	stage	run	*	weak	impImp	OffAndOn	nRp	ordProcess	imperfective
26	wy+SUF	stage	run	*	strong	impImp	OffAndOn	nRp	refProcess	imperfective
27	roz.sie	moment	enter	*	increase	impPrf	*	nRp	refProcess	perfective
28	od	moment	finish	sequential	*	impPrf	*	nRp	ordProcess	perfective
29	na	moment	finish	sequential	strong	impPrf	*	nRp	ordProcess	perfective
30	po	moment	finish	parallel	*	prfImpPrf	*	defnb	ordProcess	perfective

Hierarchy of Attributes



Conclusions

Linguistic postulates for FCA science:

1. Due to multi-valued attributes, linguistic units (viz. Concepts) exhibit multiple symmetric oppositions (structured organisations)
2. In order to represent such structurees, most often multi-dimensional attribute spaces are needed for building Contexts
3. Therefore, henceforth Conjunctive Contexts need a detailed exploration of lattice diagram representations

Linguistic postulates for cognitivists using FCA:

• Given that:
- A Context is a constitutum while Objects are its Constituentia.
- An Object is a definiendum while Attributes are its Definientia
- Definientia are justified by Explanenda (e. g.: hierarchies of features)

It is necessary to add explanatory hierarchical organisations to the collections of attributes ('definientia' in definitions need to be explained why they fit well together)

Thank you for your attention

wlodarczyk.andre@gmail.com