Introduction to Formal Concept Analysis

OntoQuery - Lecture 1

Uta Priss
School of Computing,
Napier University,
Edinburgh, UK
u.priss@napier.ac.uk

May, 2004

Concept (or Galois) Lattices were independently discovered by

- Gerard Salton (1968): document/term lattices [but his lattice retrieval models were omitted from the 2nd edition of his book!]

- Barbut & Montjardet (1970): Galois Lattices

- Yulii Shreider, Russian School of Taxonomy (1970-1980s)

- Rudolf Wille (1983): Formal Concept Analysis

- Jon Barwise & Jerry Seligman (1997): Classifications in Information Flow (relates to Chu Spaces)

Applications in many domains:

software engineering, information retrieval, classification, taxonomy, linguistics, data analysis, ontologies ....

Concept lattices can express duality

- objects/attributes
- extension/intension
- token/type
- value/data type
- data-driven/theory-driven
- bottom-up/top-down

relational algebra	concept analysis
horizontal relations	vertical relations
relational composition quantification	inheritance implications, dependencies
ER diagrams conceptual graphs	line diagrams concept lattices
computationally easy	computationally difficult

Definitions:

- set G of (formal) objects

- set M of (formal) attributes

- relation I between G and M
  gIm or $(g, m) \in I$ is read as
  'object g has attribute m'

- formal context $\mathcal{K}$ is a triple (G, M, I)

2 formal contexts:

	female	juvenile	adult	male
filly	x	x		
mare	x		x	
colt		x		x
stallion			x	x
cow	x		x	
ram			x	x
bull			x	x
ewe	x		x	
foal		x		
calf		x		
lamb		x		

	horse	cow	sheep	animal
filly	x			x
mare	x			x
colt	x			x
stallion	x			x
cow		x		x
ram			x	x
bull		x		x
ewe			x	x
foal	x			x
calf		x		x
lamb			x	x

- all common attributes of a set A of objects

$$\iota A := \{m \in M \mid gIm \text{ for all } g \in A\}$$

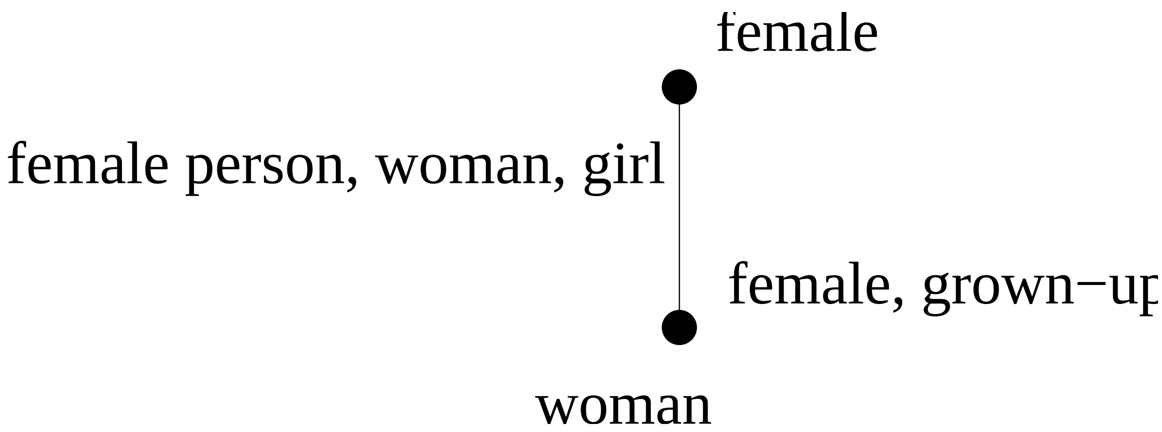
- all common objects of a set B of attributes

$$\varepsilon B := \{g \in G \mid gIm \text{ for all } m \in B\}$$

- (A, B) is a *formal concept* if

$$A \subseteq G, B \subseteq M, A = \varepsilon B \text{ and } B = \iota A$$

- A is called the *extension* ($Ext(c)$); B is called the *intension* ($Int(c)$) of a concept $c := (A, B)$



c_1 is a *subconcept* of the concept c_2 (denoted by $c_1 \leq c_2$) if

$$Ext(c_1) \subseteq Ext(c_2)$$

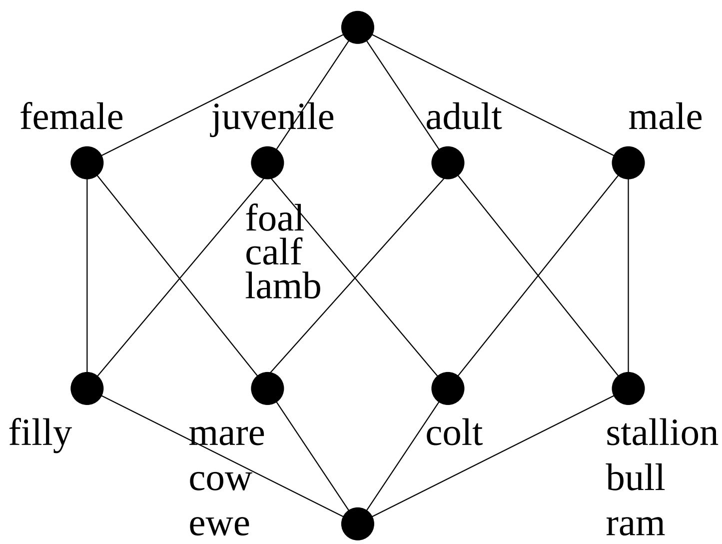
which is equivalent to

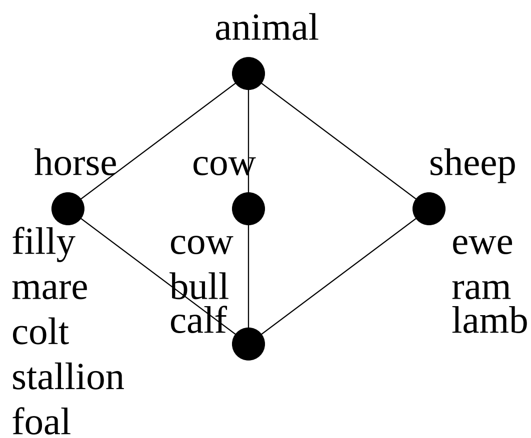
$$Int(c_2) \subseteq Int(c_1)$$

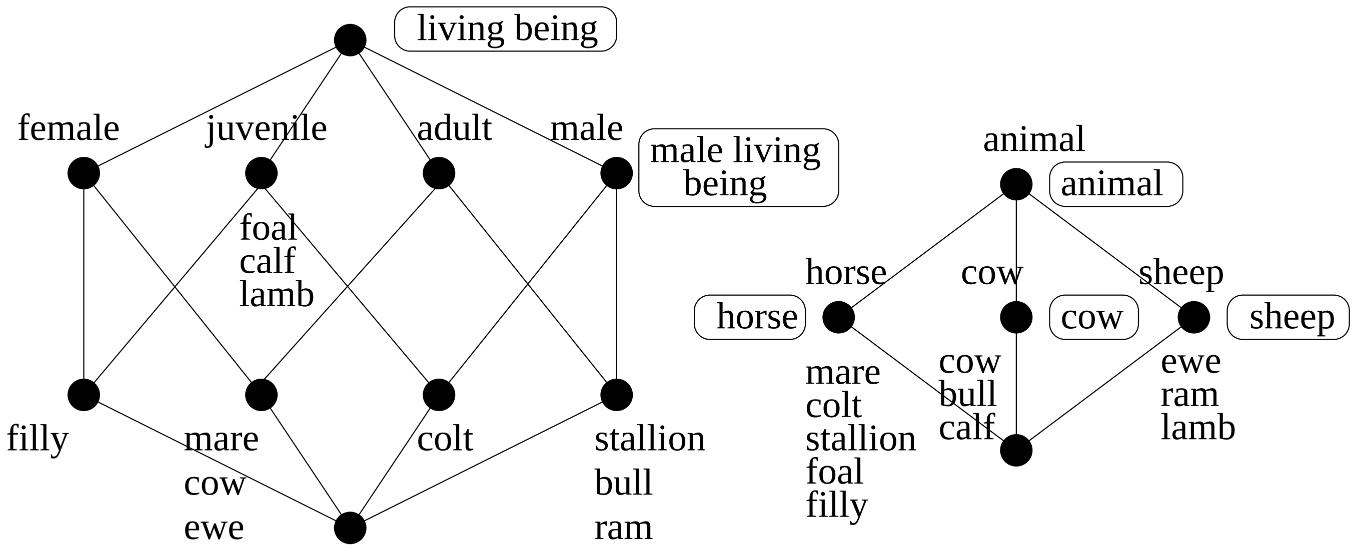
The set of all formal concepts of (G, M, I) is denoted by $\mathcal{B}(G, M, I)$. Together with the relation '$\leq$', $\mathcal{B}(G, M, I)$ forms a mathematical lattice.

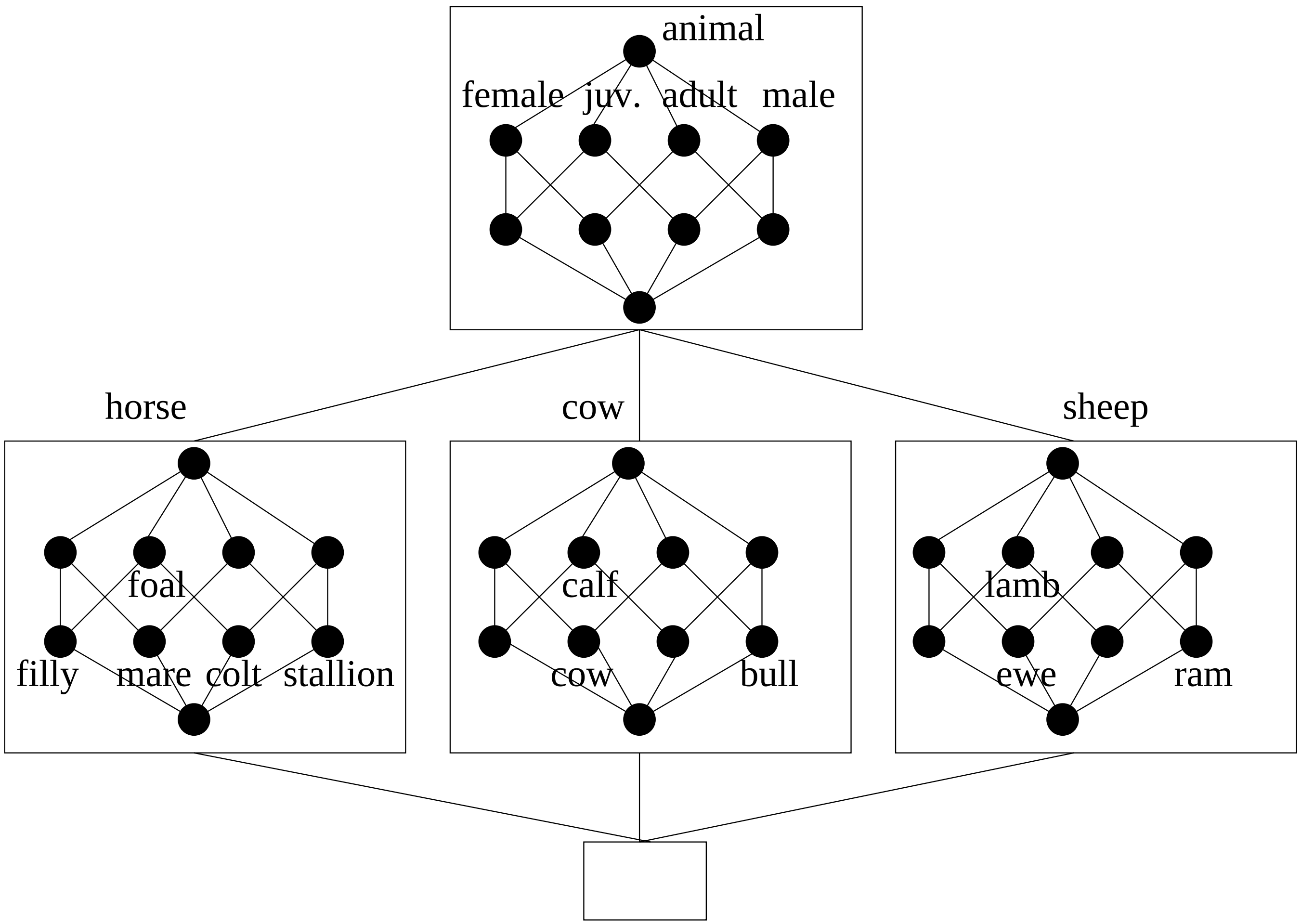
	female	juvenile	adult	male
filly	x	x		
mare	x		x	
colt		x		x
stallion			x	x
cow	x		x	
ram			x	x
bull			x	x
ewe	x		x	
foal		x		
calf		x		
lamb		x		

	horse	cow	sheep	animal
filly	x			x
mare	x			x
colt	x			x
stallion	x			x
cow		x		x
ram			x	x
bull		x		x
ewe			x	x
foal	x			x
calf		x		x
lamb			x	x






living being
female
juvenile
adult
male
male living being
foal
calf
lamb
filly
mare
cow
ewe
colt
stallion
bull
ram
animal
animal
horse
cow
sheep
horse
cow
sheep
mare
colt
stallion
foal
filly
cow
bull
calf
ewe
ram
lamb


animal
female
juv.
adult
male
horse
cow
sheep
foal
calf
lamb
filly
mare
colt
stallion
cow
bull
ewe
ram