Formal Concept Analysis

B. Ganter and Rudolf Wille. Formal Concept Analysis. Mathematical Foundations. Springer Verlag, 1999.

- Formalisation mathématique de la notion de concept et de classification conceptuelle
- Se range dans les approches symboliques de l'IA
  - classification, structuration des connaissances
  - apprentissage (règles)
  - extraction de patrons de connaissances

Data

Definition (Formal Context)

A Formal Context is a triple (O, A, R), where O is a finite set of objects, A is a finite set of attributes and $R \subseteq O \times A$ is a binary relation. $(o, a) \in R$ means that object o owns attribute a. This is also denoted by oRa.

Example (Formal Context Animals11)

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Applications associées au contexte formel

Definition (Maps of a binary relation $R \subseteq O \times A$)

Let us denote by $P(E)$ the set of subsets of a finite set E.
Map f associates to an object set the attributes they have in common.
$f : P(O) \rightarrow P(A)$
$X \longmapsto f(X) = \{y \in A \mid \forall x \in X, (x, y) \in R\}$
Map g associates to an attribute set the objets that sharing these attributes
$g : P(A) \rightarrow P(O)$
$Y \longmapsto g(Y) = \{x \in O \mid \forall y \in Y, (x, y) \in R\}$.

In Ganter and Wille, f et g are denoted by the polymorphic symbol $'$. We will use both notations depending the situations.

Applications associées au contexte formel

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

$$f(\{great-auk, silver-gull\}) = \{sea-habitat, eats-fish, water-habitat, feathered, flies\}$$

$$g(\{sea-habitat, eats-fish\}) = \{great-auk, silver-gull, greater-flamingo, little-tern, artic-tern\}$$

Clarified Formal Context

Definition (Clarified Formal Context)

A Formal Context (O, A, R) is *object−clarified* if $\forall o_1, o_2 \in O$, when $f(\{o_1\}) = f(\{o_2\})$, then $o_1 = o_2$.
A Formal Context (O, A, R) is *attribute−clarified* if $\forall a_1, a_2 \in A$, when $g(\{a_1\}) = g(\{a_2\})$, then $a_1 = a_2$.

Example (Clarified Formal Context Animals11)

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	eats-fish
ladybird	×					×			
bat	×	×							
ostrich			×						
greater-flamingo	×		×	×			×		×
silver-gull	×		×		×		×		×
great-auk	×		×				×		×
wood-pecker	×		×					×	
giant-otter									×
arctic-tern	×		×	×	×		×		×

Removal of : *little−tern* (eq. to *greater−flamingo*) ; *six−legged* (eq. to *elytra*) ; *water−habitat* (eq. to *eats−fish*)

Reduced Formal Context

Definition (Reduced Formal Context)

A Formal Context (O, A, R) is *object−reduced* if it does not contain reducible object (all objects are said *irreducible*), i.e., $\forall o \in O,\ \nexists X \subseteq O, o \notin X$ with $f(\{o\}) = f(X)$.

A Formal Context (O, A, R) is *attribute−reduced* if it does not contain reducible attributes (all attributes are said *irreducible*), i.e., $\forall a \in A,\ \nexists Y \subseteq A, a \notin Y$ with $g(\{a\}) = g(Y)$.

Example (Reduced and clarified Formal Context Animals11)

	flies	nocturnal	feathered	migratory	red-bill	elytra	wood-habitat	eats-fish
ladybird	×					×		
bat	×	×						
ostrich			×					
greater-flamingo	×		×	×				×
silver-gull	×		×		×			×
wood-pecker	×		×				×	
giant-otter								×
arctic-tern	×		×	×	×			×

Removal of :

great−auk row = *greater−flamingo* row ∩ *silver−gull* row.

sea−habitat column = *eats−fish* column ∩ *feathered* column ∩ *flies* column.

Formal Concept

Definition (Formal Concept)

A Formal Concept C of a formal context (O, A, R) is a pair $C{=}(E,\ I)$ such that $f(E) = I$ (or equivalently $E = g(I)$). E is the concept *extent* ; I is the concept *intent*.

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Example (Example)

The following set pairs is a formal concept $C_{great-auk}$:
$X_3 = \{great-auk, silver-gull, greater-flamingo, little-tern, arctic-tern\}$
$Y_3 = \{sea-habitat, eats-fish, water-habitat, feathered, flies\}$
This concept groups flying birds (feathered, flies) which leave close to the sea, and eat fishes.

Formal Concept

Definition (Formal Concept)

A Formal Concept C of a formal context (O, A, R) is a pair $C{=}(E,\ I)$ such that $f(E) = I$ (or equivalently $E = g(I)$). E is the concept *extent* ; I is the concept *intent*.

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Example (Counter-example of animal concepts)

The following set pairs does not form a formal concept :
$X_1 = \{great{-}auk, silver{-}gull\}$
$Y_1 = \{sea{-}habitat, eats{-}fish, water{-}habitat, feathered, flies\}$
because other objects have the I_1 attributes (the object set if not maximal) :
greater–flamingo, *little–tern*, *arctic–tern*

Formal Concept

Definition (Formal Concept)

A Formal Concept C of a formal context (O, A, R) is a pair $C{=}(E,\ I)$ such that $f(E) = I$ (or equivalently $E = g(I)$). E is the concept *extent* ; I is the concept *intent*.

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Example (Counter-example of animal concepts)

The following set pairs does not form a formal concept :
$X_2 = \{great{-}auk, silver{-}gull, greater{-}flamingo, little{-}tern, arctic{-}tern\}$
$Y_2 = \{sea{-}habitat, eats{-}fish\}$
because other attributes are common to the E_2 objects (the attribute set if not maximal) : *water–habitat*, *feathered*, *flies*

Concept ordering

Definition (Concept ordering)

Concepts can be ordered through the following partial order $\leq_s$:
$(E_1, I_1) \leq_s (E_2, I_2) \Leftrightarrow E_1 \subseteq E_2$
(or equivalently $I_2 \subseteq I_1$)
(E_1, I_1) is called a sub-concept of (E_2, I_2), (E_2, I_2) is called a super-concept of (E_1, I_1).

Example (Example of concept ordering)

$C_{great-auk}$:
$X_3 = \{great-auk, silver-gull, greater-flamingo, little-tern, arctic-tern\}$
$Y_3 = \{sea-habitat, eats-fish, water-habitat, feathered, flies\}$

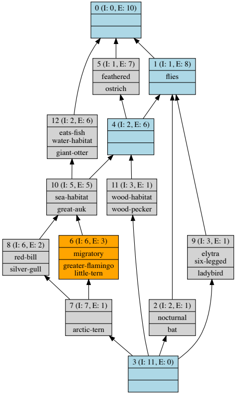
$C_{silver-gull}$:
$X_4 = \{silver-gull, arctic-tern\}$
$Y_4 = \{sea-habitat, eats-fish, water-habitat, feathered, flies, red-bill\}$

$C_{silver-gull} \leq_s C_{great-auk}$, as $X_4 \subseteq X_3$ and $Y_3 \subseteq Y_4$.

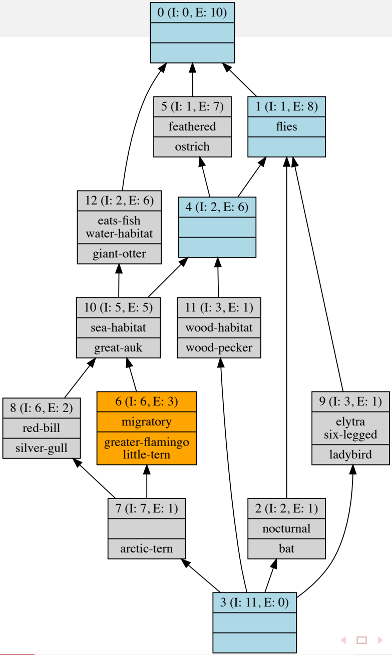
Concept lattice

Definition (Concept lattice)

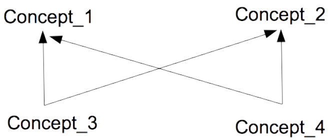
The set $\mathcal{C}_K$ of all concepts of a formal context $K = (O, A, R)$, provided with the partial order $\leq_s$ is called the concept lattice associated with K. It is denoted by $\mathcal{L}_K = (\mathcal{C}_K, \leq_s)$.



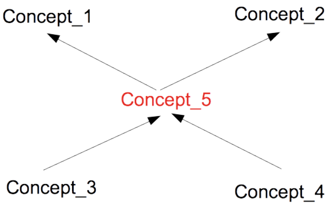
Concept lattice



Property of the lattice structure



Partial order



Partial order with lattice structure

The partial order of the left-hand-side is not a lattice : $Concept\_3$ and $Concept\_4$ have two more specific superconcepts ($Concept\_1$ and $Concept\_2$) ; $Concept\_1$ and $Concept\_2$ have two more general subconcepts ($Concept\_3$ and $Concept\_4$).

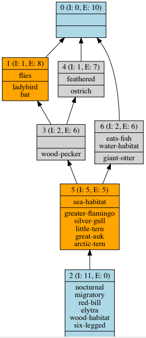
The partial order of the right-hand-side is a lattice : $Concept\_3$ and $Concept\_4$ have a unique most specific superconcept (upper bound $Concept\_5$) ; $Concept\_1$ and $Concept\_2$ have a unique most general subconcept (lower bound $Concept\_5$). This is denoted in literature : $Concept\_5 = Concept\_3 \vee Concept\_4$ and $Concept\_5 = Concept\_1 \wedge Concept\_2$.

Iceberg lattice for minimal support n

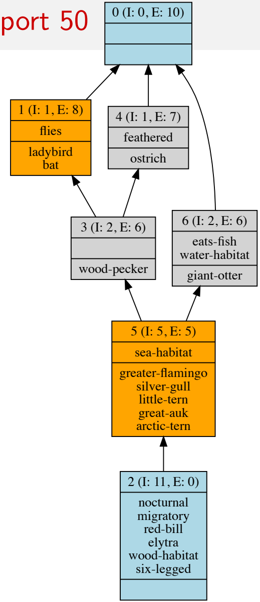
Definition (Iceberg lattice for minimal support n)

For a formal context $K = (O, A, R)$, let us consider the set $\mathcal{C}_{M_K}$, composed of the bottom concept of $\mathcal{L}_K$ and all the concepts $C = (E, I)$ of K , such that $|E| \geq \frac{n \times |O|}{100}$. $(\mathcal{C}_{M_K}, \leq_s)$ is called the Iceberg lattice associated with K, for minimal support n. It is denoted by $\mathcal{I}CEBERG_{K_n} = (\mathcal{C}_{M_K}, \leq_s)$.

Each concept has at least 5 objects in its extent (except the bottom)



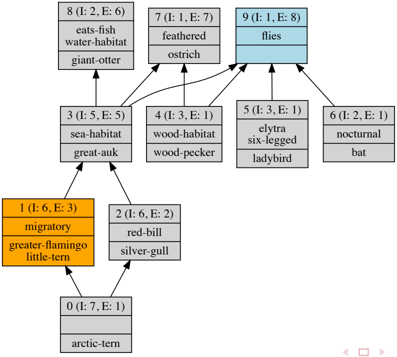
Iceberg lattice for min. support 50



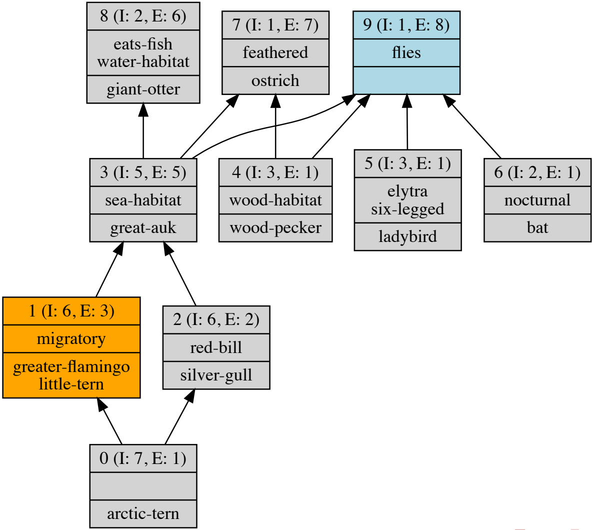
AOC-poset

Definition (AOC-poset)

The set $\mathcal{C}_{I_K}$ of all introducer concepts (concepts that introduce an object, or an attribute, or both) of a formal context $K = (O, A, R)$, provided with the partial order $\leq_s$ is called the AOC-poset (Attribute-Object-Concept partially ordered set) associated with K. It is denoted by $\mathcal{AOC}_K = (\mathcal{C}_{I_K}, \leq_s)$.



AOC-poset


8 (I: 2, E: 6)
eats-fish
water-habitat
giant-otter
7 (I: 1, E: 7)
feathered
ostrich
9 (I: 1, E: 8)
flies
3 (I: 5, E: 5)
sea-habitat
great-auk
4 (I: 3, E: 1)
wood-habitat
wood-pecker
5 (I: 3, E: 1)
elytra
six-legged
ladybird
6 (I: 2, E: 1)
nocturnal
bat
1 (I: 6, E: 3)
migratory
greater-flamingo
little-tern
2 (I: 6, E: 2)
red-bill
silver-gull
0 (I: 7, E: 1)
arctic-tern

Size of the conceptual structures

For a formal context $K = (O, A, R)$,

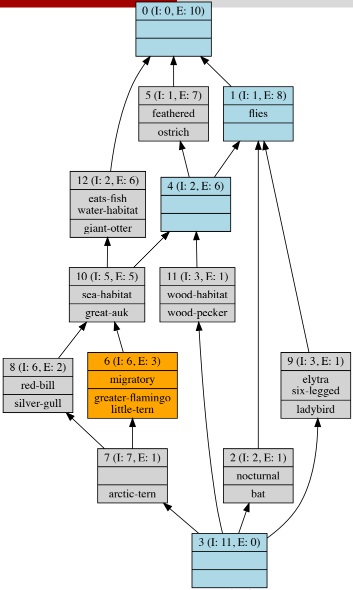
- The concept lattice may have up to $2^{min(|A|,|O|)}$ concepts.
  This extreme situation is reached with the lattice of all subsets of E, where E is either O if $|O| = min(|A|, |O|)$, or A if $|A| = min(|A|, |O|)$.
- The AOC-poset may have up to $|A| + |O|$ concepts, since a concept introduces either an object or an attribute.
  This bound is reached for example when $|A| = |O|$ and every attribute is shared by several distinct objects (with a bipartite crown graph for example).
- Despite this difference, there are formal contexts where the concept lattice and the AOC-poset are identical, when every concept of the lattice is an introducer.

Reducible / Irreducible elements

Definition (Irreducible/Reducible concept)

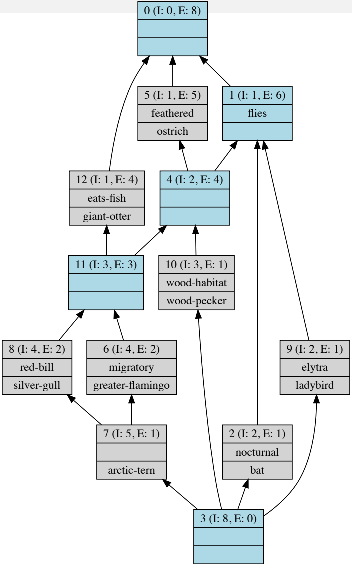
Let us consider a conceptual structure.
A concept is *sup-reducible* if it is the upper bound of several other concepts.
The bottom is *sup-reducible*, as it is considered the upper bound of $\emptyset$.

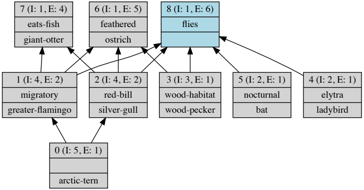
A concept is *inf-reducible* if it is the lower bound of several other concepts.
The top is *inf-reducible*, as it is considered the lower bound of $\emptyset$.



*Concept*_10 is sup-reducible, because it is the upper bound of *Concept*_8 and *Concept*_6.
*Concept*_10 is inf-reducible, because it is the lower bound of *Concept*_4 and *Concept*_12.
*Concept*_1 is sup-reducible, because it is the upper bound of *Concept*_4, *Concept*_2 and *Concept*_9, but it is inf-irreducible, because it is not a lower bound of several other concepts.

AOC-poset and irreducible elements





For a reduced formal context $K = (O, A, R)$, the AOC-poset is the set of irreducible concepts of the lattice.

Formal Context and propositional logic

Definition (Implication)

An implication of a formal context $K = (O, A, R)$ with associated maps f, g, denoted by $Prem \implies Conc$, is a pair of attribute sets $(Prem, Conc)$, with $Prem, Conc \subseteq A$ where all the objects that own the attributes of $Prem$ (premise) also own the ones of B (conclusion) : $g(Prem) \subseteq g(Conc)$, or equivalently $\{o|\forall a_{prem} \in Prem, (o, a_{prem}) \in R\} \subseteq \{o|\forall a_{conc} \in Conc, (o, a_{conc}) \in R\}$.

Binary implications of a formal context

Minimal non-redundant set of binary implications for animals :

- *migratory* $\implies$ *sea-habitat*, from $Concept\_1 \leq_s Concept\_3$
- *red-bill* $\implies$ *sea-habitat*, from $Concept\_2 \leq_s Concept\_3$
- *sea-habitat* $\implies$ *eats-fish*, from $Concept\_3 \leq_s Concept\_8$
- *eats-fish* $\implies$ *water-habitat* and *water-habitat* $\implies$ *eats-fish*, from $Concept\_8$ which introduces both
- *sea-habitat* $\implies$ *feathered*, from $Concept\_3 \leq_s Concept\_7$
- *sea-habitat* $\implies$ *flies*, from $Concept\_3 \leq_s Concept\_9$
- *wood-habitat* $\implies$ *flies*, from $Concept\_4 \leq_s Concept\_9$
- *elytra* $\implies$ *flies*, from $Concept\_5 \leq_s Concept\_9$
- *elytra* $\implies$ *six-legged* and *six-legged* $\implies$ *elytra*, from $Concept\_5$ which introduces both
- *nocturnal* $\implies$ *flies*, from $Concept\_6 \leq_s Concept\_9$

Duquenne-Guigues Basis of Implications (DGBI)

Cardinality minimal set of non redundant implications

<0> flies,feathered,sea-habitat,wood-habitat,eats-fish,water-habitat => nocturnal,migratory,red-bill,elytra,six-legged

<0> flies,feathered,elytra,six-legged => nocturnal,migratory,red-bill,sea-habitat,wood-habitat,eats-fish,water-habitat

<0> flies,nocturnal,elytra,six-legged => feathered,migratory,red-bill,sea-habitat,wood-habitat,eats-fish,water-habitat

<0> flies,nocturnal,feathered => migratory,red-bill,elytra,sea-habitat,wood-habitat,six-legged,eats-fish,water-habitat

<1> six-legged => flies,elytra

<1> wood-habitat => flies,feathered

<1> elytra => flies,six-legged

<1> nocturnal => flies

<2> red-bill => flies,feathered,sea-habitat,eats-fish,water-habitat

<3> migratory => flies,feathered,sea-habitat,eats-fish,water-habitat

<5> sea-habitat => flies,feathered,eats-fish,water-habitat

<5> feathered,eats-fish,water-habitat => flies,sea-habitat

<5> flies,eats-fish,water-habitat => feathered,sea-habitat

<6> water-habitat => eats-fish

<6> eats-fish => water-habitat

FCA in Artificial Intelligence

- Unsupervised/Supervised versions
- Robustness
- Symbolic machine learning
- Hierarchical classification (multiple)
- Knowledge Navigation
- Explanation
- Generality

Application

- Construire les structures du cours avec FCA4J https://www.lirmm.fr/fca4j/
- Se familiariser avec RCAviz https://info-demo.lirmm.fr/rcaviz/
- Fichiers et liens disponibles sur Moodle https://moodle.umontpellier.fr/course/view.php?id=22617