Formale Begriffsanalyse

Formale Begriffsanalyse (Concept Analysis)

- Lernziele
  - Einführung von Begriffsanalyse
- Kontext
  - Begriffsanalyse ist Technik zur Analyse binärer Relationen
  - Begriffsanalyse wird vielseitig eingesetzt; Beispiele:
    - zur Lokalisierung von Merkmalen
    - zur Restrukturierung von Vererbungshierarchien

Von Objekten und Attributen zu Klassen

	four-legged	hair-covered	intelligent	marine	thumbed
cats	×	×			
dogs	×	×			
dolphins			×	×	
gibbons		×	×		×
humans			×		×
whales			×	×	

Begriffsanalyse (Concept Analysis) I

- Menge $\mathcal{O}$ von Objekten
- Menge $\mathcal{A}$ von Attributen
- Relation $\mathcal{I} \subseteq \mathcal{O} \times \mathcal{A}$
- das Tripel $K = (\mathcal{O}, \mathcal{A}, \mathcal{I})$ wird formaler Kontext genannt
- für eine Menge von Objekten $O \subseteq \mathcal{O}$ ist $\sigma(O)$ die Menge gemeinsamer Attribute:

$$\sigma(O) := \{a \in \mathcal{A} \quad | \quad (o, a) \in \mathcal{I} \text{ für alle } o \in O\}$$

- für eine Menge von Attributen $A \subseteq \mathcal{A}$ ist $\tau(A)$ die Menge der gemeinsamen Objekte:

$$\tau(A) := \{o \in \mathcal{O} \quad | \quad (o, a) \in \mathcal{I} \text{ für alle } a \in A\}$$

Begriffsanalyse (Concept Analysis) II

- Ein Paar aus Objekten und Attributen $b = (O, A)$ heißt **Begriff (Concept)**, genau dann, wenn

$$A = \sigma(O)$$

und gleichzeitig

$$O = \tau(A)$$

- ein Begriff entspricht einem maximalen Rechteck in der Tabelle (modulo Zeilen- und Spaltenpermutationen)
- für einen Begriff $b = (O, A)$ ist
  - $O = \mathit{Umfang}(b)$ (extent) und
  - $A = \mathit{Inhalt}(b)$ (intent) des Begriffs b.

	four-legged	hair-covered	intelligent	marine	thumbed
cats	×	×			
dogs	×	×			
dolphins			×	×	
gibbons		×	×		×
humans			×		×
whales			×	×	

({dogs, cats}, {hair-covered, four-legged})
({humans, gibbons}, {thumbed, intelligent})
({gibbons}, {thumbed, intelligent hair-covered})
({whales, humans, gibbons, dolphins}, {intelligent})
({gibbons, dogs, cats}, {hair-covered})
({whales, dolphins}, {marine, intelligent})
({whales, humans, gibbons, dolphins, dogs, cats}, {})
({}, {thumbed, marine, intelligent, hair-covered, four-legged})

Partielle Ordnung der Begriffe

hierarchische Ordnung $\leq$: $b_1 = (O_1, A_1)$ und $b_2 = (O_2, A_2)$ zwei Begriffe des selben formalen Kontexts, dann ist

$$b_1 \leq b_2 \Leftrightarrow O_1 \subseteq O_2$$

oder, dual dazu,

$$b_1 \leq b_2 \Leftrightarrow A_2 \subseteq A_1$$

- $b_1 \leq b_2$: b_2 **Oberbegriff (Superconcept)** zu b_1
- b_1 als **Unterbegriff (Subconcept)** zu b_2

$\Rightarrow$ b_2 hat mindestens so viele Objekte wie b_1 bzw. b_1 hat mindestens so viele Attribute wie b_2

- ({gibbons}, {thumbed, intelligent hair-covered}) $\leq$ ({humans, gibbons}, {thumbed, intelligent})
- Falls weder $b_1 \leq b_2$ noch $b_2 \leq b_1$, dann sind b_1 und b_2 **unvergleichbar**

Begriffsverband

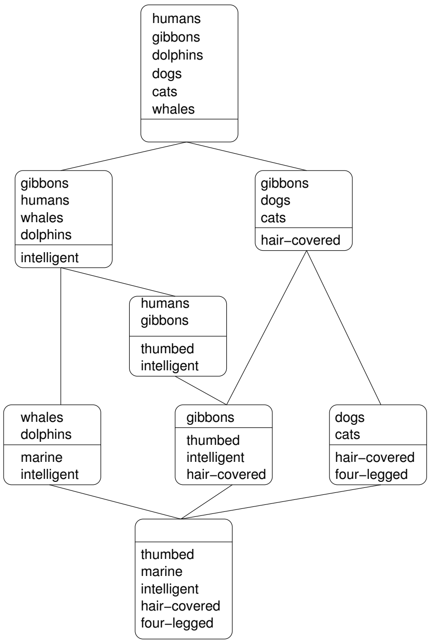
Menge $\mathcal{B}$ aller Begriffe eines Kontexts K zusammen mit Halbordnung $\leq$ bilden einen vollständigen Verband, den so genannten **Begriffsverband**:

$$\mathcal{B}(K) = \{(O, A) \in 2^{\mathcal{O}} \times 2^{\mathcal{A}} \quad | \quad A = \sigma(O) \text{ und } O = \tau(A)\}$$

Hasse-Diagramme

Hasse-Diagramme visualisieren die Relation $<$:

$b_1 < b_2 \Leftrightarrow b_1 \leq b_2$ und es gibt keinen Begriff $b(\neq b_1, b_2)$, mit $b_1 \leq b \leq b_2$


humans
gibbons
dolphins
dogs
cats
whales
gibbons
humans
whales
dolphins
intelligent
gibbons
dogs
cats
hair−covered
humans
gibbons
thumbed
intelligent
whales
dolphins
marine
intelligent
gibbons
thumbed
intelligent
hair−covered
dogs
cats
hair−covered
four−legged
thumbed
marine
intelligent
hair−covered
four−legged

Infimum

Für zwei Begriffe b_1 und b_2

- **Gemeinsamer Unterbegriff Infimum** $\wedge$

$$(O_1, A_1) \wedge (O_2, A_2) = (O_1 \cap O_2, \sigma(O_1 \cap O_2))$$

Begriff, der die Menge der gemeinsamen Attribute zweier Objektmengen enthält

Infimum

Für zwei Begriffe b_1 und b_2

- **Gemeinsamer Unterbegriff Infimum** $\wedge$

$$(O_1, A_1) \wedge (O_2, A_2) = (O_1 \cap O_2, \sigma(O_1 \cap O_2))$$

Begriff, der die Menge der gemeinsamen Attribute zweier Objektmengen enthält

$$\begin{aligned}
&(\{human, gibbons\}, \{thumbed, intelligent\}) \\
&\wedge(\{gibbons, dogs, cats\}, \{hair\text{-}covered\}) \\
&= (\{human, gibbons\} \cap \{gibbons, dogs, cats\}, \\
&\sigma(\{human, gibbons\} \cap \{gibbons, dogs, cats\})) \\
&= (\{gibbons\}, \{thumbed, intelligent, hair\text{-}covered\})
\end{aligned}$$

Supremum

Für zwei Begriffe b_1 und b_2

- **Gemeinsamer Oberbegriff Supremum ($\vee$)**

$$(O_1, A_1) \vee (O_2, A_2) = (\tau(A_1 \cap A_2), A_1 \cap A_2)$$

Begriff, der die Menge der gemeinsamen Objekte zweier Attributmengen umfasst

Supremum

Für zwei Begriffe b_1 und b_2

- **Gemeinsamer Oberbegriff Supremum ($\vee$)**

$$(O_1, A_1) \vee (O_2, A_2) = (\tau(A_1 \cap A_2), A_1 \cap A_2)$$

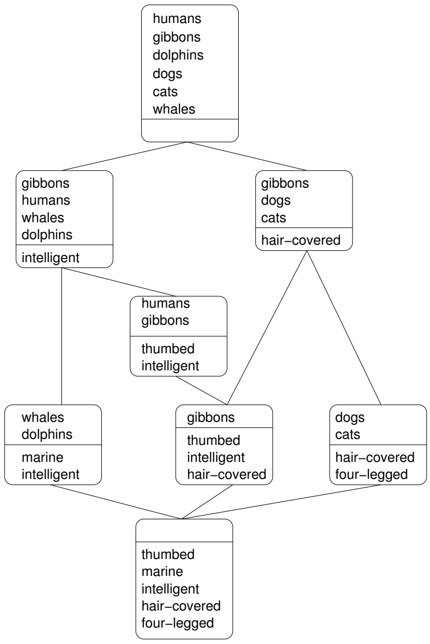
Begriff, der die Menge der gemeinsamen Objekte zweier Attributmengen umfasst

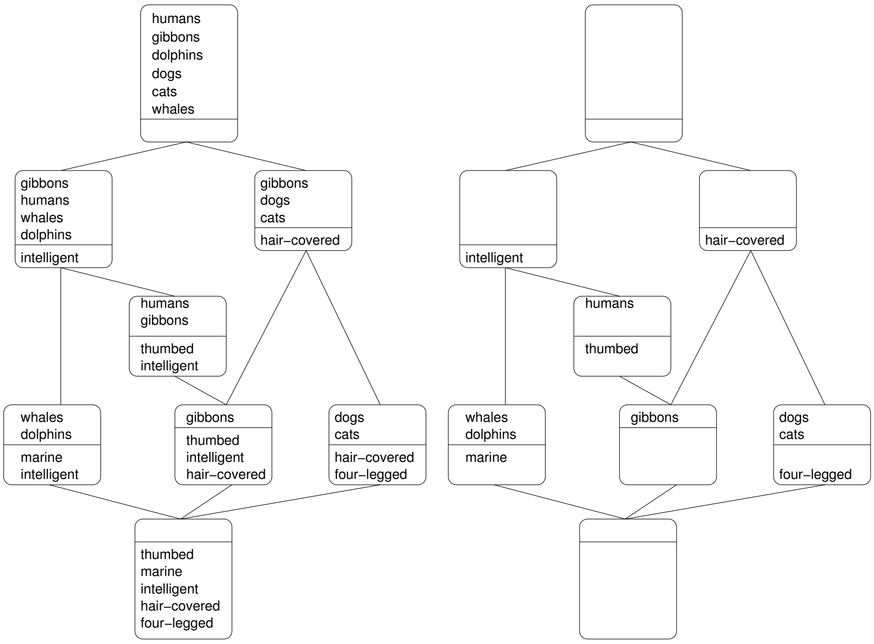
$$\begin{aligned} & (\{gibbons\}, \{thumbed, intelligent, hair\text{-}covered\}) \\ & \vee(\{dogs, cats\}, \{four\text{-}legged, hair\text{-}covered\}) \\ = & (\tau(\{thumbed, intelligent, hair\text{-}covered\} \cap \{four\text{-}legged, hair\text{-}covered\}), \\ & \{thumbed, intelligent, hair\text{-}covered\} \cap \{four\text{-}legged, hair\text{-}covered\}) \\ = & (\{gibbons, dogs, cats\}, \{hair\text{-}covered\}) \end{aligned}$$

Begriffsverband

Für die Operatoren $\wedge$ und $\vee$ gelten:

- Kommutativität: $x \vee y = y \vee x$ und $x \wedge y = y \wedge x$
- Assoziativität: $(x \vee y) \vee z = x \vee (y \vee z)$ und $(x \wedge y) \wedge z = x \wedge (y \wedge z)$
- Absorbtionsgesetze (Verschmelzungsgesetze): $x \vee (x \wedge y) = x$ und $x \wedge (x \vee y) = x$
- Größtes (Top, 1) und kleinstes (Bottom, 0) Element existieren


humans
gibbons
dolphins
dogs
cats
whales
gibbons
humans
whales
dolphins
intelligent
gibbons
dogs
cats
hair–covered
humans
gibbons
thumbed
intelligent
whales
dolphins
marine
intelligent
gibbons
thumbed
intelligent
hair–covered
dogs
cats
hair–covered
four–legged
thumbed
marine
intelligent
hair–covered
four–legged


humans
gibbons
dolphins
dogs
cats
whales
gibbons
humans
whales
dolphins
intelligent
gibbons
dogs
cats
hair–covered
humans
gibbons
thumbed
intelligent
whales
dolphins
marine
intelligent
gibbons
thumbed
intelligent
hair–covered
dogs
cats
hair–covered
four–legged
thumbed
marine
intelligent
hair–covered
four–legged
intelligent
hair–covered
humans
thumbed
whales
dolphins
marine
gibbons
dogs
cats
four–legged

Verkürzter Verband

Eindeutig bestimmter Knoten im Begriffsverband, der mit

- Attribut a beschriftet wird:

  $$\mu(a) = \bigvee\{b \in \mathcal{B}(K) \quad | \quad a \in \mathit{Inhalt}(b)\}$$

  heißt **Attributbegriff** zu a ($\rightarrow$ größter Begriff mit a)
- Der Knoten, der mit dem Objekt o markiert wird:

  $$\gamma(o) = \bigwedge\{b \in \mathcal{B}(K) \quad | \quad o \in \mathit{Umfang}(b)\}$$

  und heißt **Objektbegriff** zu o ($\rightarrow$ kleinster Begriff mit o).

Von Objekten und Attributen zu Klassenhierarchien

- Erstelle Begriffsverband für Tabelle.
- Jedes Konzept ist prinzipiell eine Klasse.
- Die $\leq$-Relation stellt die Vererbung dar.
- Supremum: gemeinsame Oberklasse.
- Infimum: gemeinsame Unterklasse.

Literaturreferenzen

1 Bernhard Ganter and Rudolf Wille. Formale Begriffsanalyse: mathematische Grundlagen Springer Verlag, 1996.