

Formale Begriffsanalyse (Concept Analysis)

f

- Lernziele
 - Verständnis von Begriffsanalyse
 - Begriffsanalyse anwenden können
- 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)

- 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

Begriffsanalyse (Concept Analysis)

- 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**:

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

Begriffsanalyse (Concept Analysis)

- 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**:

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

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

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

Begriffsanalyse (Concept Analysis)

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

$$A = \sigma(O)$$

und gleichzeitig

$$O = \tau(A)$$

Begriffsanalyse (Concept Analysis)

- 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)

Begriffsanalyse (Concept Analysis)

- 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
 - $O = \text{Umfang}(b)$ (extent) und
 - $A = \text{Inhalt}(b)$ (intent) des Begriffs b .

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

$$(\mathcal{O}, A)$$

↑

$$(\{\text{cats}, \dots, \text{whales}\}, \emptyset)$$

$$(\emptyset, \{\text{four-legged}, \dots, \text{thumbed}\})$$

	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,

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

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

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 ist **Oberbegriff (Superconcept)** zu b_1
 - b_1 ist **Unterbegriff (Subconcept)** zu b_2
- $\Rightarrow b_2$ hat mindestens die Objekte von b_1 bzw. b_1 hat mindestens die Attribute von b_2
- Beispiel: $(\{\text{gibbons}\}, \{\text{thumbed}, \text{intelligent hair-covered}\})$
 $\leq (\{\text{humans}, \text{gibbons}\}, \{\text{thumbed}, \text{intelligent}\})$

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,

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

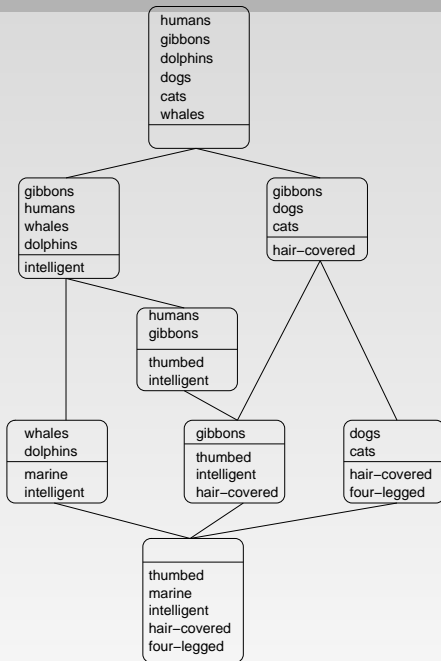
- $b_1 \leq b_2$: b_2 ist **Oberbegriff (Superconcept)** zu b_1
 - b_1 ist **Unterbegriff (Subconcept)** zu b_2
- \Rightarrow b_2 hat mindestens die Objekte von b_1 bzw. b_1 hat mindestens die Attribute von b_2
- Beispiel: $(\{\text{gibbons}\}, \{\text{thumbed}, \text{intelligent hair-covered}\})$
 $\leq (\{\text{humans}, \text{gibbons}\}, \{\text{thumbed}, \text{intelligent}\})$
 - Falls weder $b_1 \leq b_2$ noch $b_2 \leq b_1$, dann sind b_1 und b_2 **unvergleichbar**

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

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

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 $b_1 \leq b \leq b_2$



Infimum

Für zwei Begriffe b_1 und b_2

- **Gemeinsamer Unterbegriff Infimum \wedge**

$$(\textcolor{blue}{O}_1, A_1) \wedge (\textcolor{red}{O}_2, A_2) = (\textcolor{blue}{O}_1 \cap \textcolor{red}{O}_2, \sigma(\textcolor{blue}{O}_1 \cap \textcolor{red}{O}_2))$$

Begriff, der die Menge der gemeinsamen Attribute zweier Objektmengen enthält

$$\begin{array}{ccc} (O_1, A_1) & & (O_2, A_2) \\ & \searrow \quad \swarrow & \\ & (O_1 \cap O_2, A_1 \cup A_2 \cup \{...\}) & \end{array}$$

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

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

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

$$\begin{array}{ccc} & (O_1 \cup O_2 \cup \{\dots\}, A_1 \cap A_2) & \\ \swarrow & & \searrow \\ (O_1, A_1) & & (O_2, A_2) \end{array}$$

Supremum

Für zwei Begriffe b_1 und b_2

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

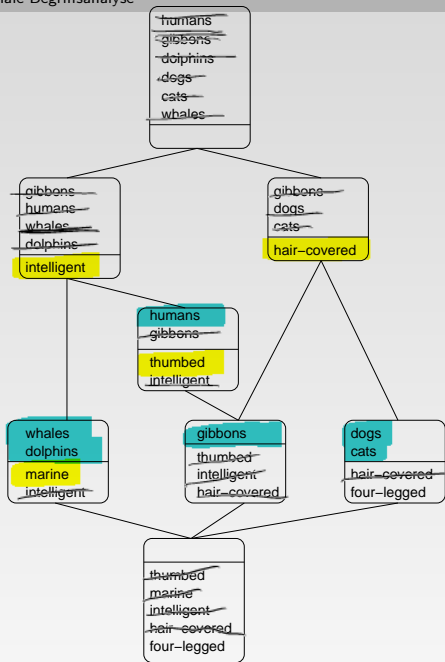
$$(O_1, A_1) \vee (O_2, A_2) = (\tau(A_1 \cap A_2), A_1 \cap A_2)$$

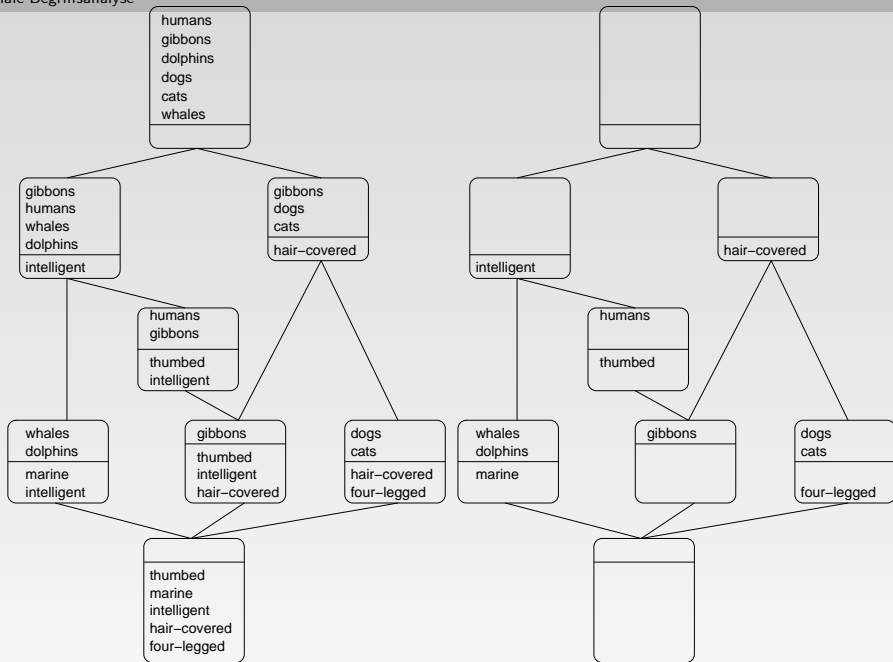
Begriff, der die Menge der gemeinsamen Objekte zweier Attributmengen umfasst

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

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





Eindeutig bestimmter Knoten im Begriffsverband, der mit

- Attribut a beschriftet wird:

$$\mu(a) = \bigvee \{b \in \mathcal{B}(K) \mid a \in \text{Inhalt}(b)\}$$

heißt **Attributbegriff** zu a (\rightarrow größter Begriff mit a)

- Der Knoten, der mit dem Objekt o markiert wird:

$$\gamma(o) = \bigwedge \{b \in \mathcal{B}(K) \mid o \in \text{Umfang}(b)\}$$

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.

- 2 Merkmalsuche
 - Ausgangsszenario für Merkmalsuche
 - Statische Analyse
 - Dynamische Analyse
 - Probleme der dynamischen Analyse
 - Statische und dynamische Analysen
 - Fallstudien
 - Inkrementelle Analyse
 - Szenarien, Merkmale, Routinen
 - Bewertung

- Lernziele
 - Ausnutzung aller verfügbarer Information (statisch und dynamisch)
 - Suche nach Merkmalen (statt Programmmustern)
- Kontext
 - Weitere Anwendung der Begriffsanalyse
 - Unterstützung der Lokalisierung beim Programmverstehen

Szenario für Merkmalsuche

- Ihre Aufgabe ist es, für ein System mit mehr als 10,000 Routinen
 - ein Merkmal hinzuzufügen, zu korrigieren oder zu ändern
 - oder einfach die Implementierung eines Merkmals zu verstehen.
- Wo beginnen Sie?
- Merkmalsuche:
 - welche Komponenten eines Systems implementieren ein einzelnes Merkmal bzw. eine Menge von Merkmalen?

Merkmale und Komponenten

Ein **Merkmal** ist eine realisierte funktionale oder nicht-funktionale Eigenschaft des Systems.

Eine **Komponente** ist eine ausführbare Einheit, z.B.:

- eine einzelne Anweisung oder ein Ausdruck,
- eine Routine (Funktion, Prozedur),
- ein Modul,
- ein Subsystem,
- eine Task, ein Thread oder Prozess.

Im Folgenden:

- Merkmal = von außen anstoßbares und beobachtbares Verhalten des Systems, das auf Komponenten abgebildet werden kann.
- Komponente = Routine.