Wissensrepräsentation
–
Christopher Habel, Hedda Schmidtke
Sommersemester 2004

Sitzung 7: Beschreibungslogiken
- Grundidee
- Ein Beispiel: $\mathcal{DL}$ (B & L, 2001)
- Semantik und Verarbeitung

Literatur: Brachman & Levesque, 2001, Kapitel 9

Objektrepräsentation: Strukturierte Beschreibungen

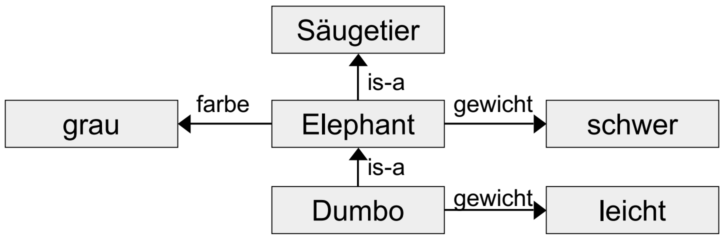
Objekte
- gehören zu Kategorien (= Mengen von Objekten)
- erfüllen Beschreibungen (= sprachliche Ausdrücke)
- instanziieren Konzepte (= mentale Repräsentationen von Klassen)

Konzepte
- entsprechen der (kognitiven) Bedeutung von Beschreibungen
- haben Kategorien als Extension

Beschreibungslogiken
- Format für komplexe Objektbeschreibungen
- Beschreibungen werden auch als ‚Konzept' bezeichnet

Historischer Ursprung von Beschreibungslogiken

Semantische Netze
- Repräsentationen für Konzept-Wissen
- Graphen mit Knoten für Konzepte und Individuen und beschrifteten Pfeilen für binäre Relationen
- Definiert über Bilder, keine klare Semantik, keine spezifizierten Verarbeitungsmechanismen



Beschreibungslogik-Ansatz

Im Kontrast zu semantischen Netzen
- Logik
  - Explizite modelltheoretische Semantik
  - Klare Syntax und Korrektheit von Inferenzverfahren
- Unterscheidung von
  - is-a zwischen Konzepten: Subsumption; und
  - is-a zwischen Objekten und Konzepten: Instanziierung
  - Diese beiden Relationen gehören zu den ausgezeichneten Basisbausteinen mit fester Interpretation
- Fokus auf Verarbeitungsmöglichkeiten und Berechenbarkeitsfragen

Konzeptsysteme, Systeme von Beschreibungen

Konzepte
- können allgemeiner / spezifischer sein als andere Konzepte (Subsumption)
- können einander ausschließen (Exklusion)
- können durch eine atomare oder komplexe Beschreibung gegeben sein

Objekte
- können in Relation zueinander stehen

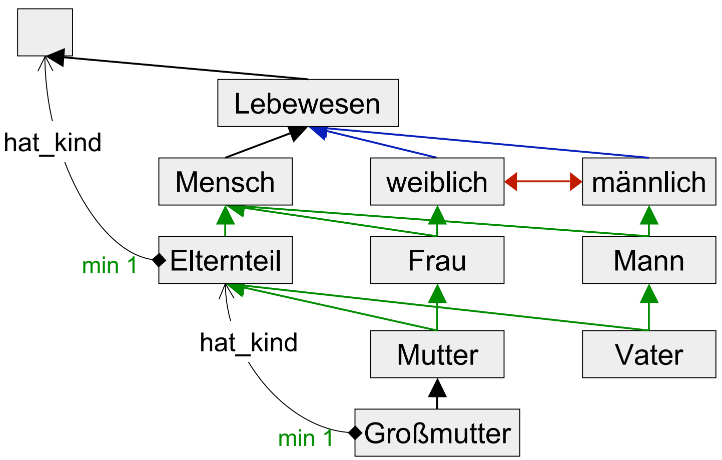
Konzepte
- können Informationen über Relationen beinhalten

Objektbeschreibungen
- in natürlicher Sprache: Nominalphrasen

Beispiel: Konzeptsysteme (sprachlich)

‚Lebewesen' ist ein Konzept.
‚Mensch' ist Subkonzept von ‚Lebewesen'.
‚weiblich' ist Subkonzept von ‚Lebewesen'.
‚Frau' steht für ‚weiblicher Mensch'.
‚männlich' steht für ‚nicht-weibliches Lebewesen'.
‚Mann' steht für ‚männlicher Mensch'.
‚hat_kind' ist eine Relation.
‚Elternteil' steht für ‚Mensch, der (mind.) ein Kind hat'.
‚Mutter' steht für ‚weiblicher Elternteil'.
‚Vater' steht für ‚männlicher Elternteil'.
‚Großmutter' steht für ‚Frau, die (mind.) ein Kind hat, das ein Elternteil ist'.

Beispiel: Konzeptsysteme (graphisch)



Konzepte / Beschreibungen

Primitive Konzepte / Atomare Beschreibungen
- keine Definition vorhanden
- nur notwendige Bedingungen bekannt
- explizite Einordnung in Subsumptionshierarchie

Definierte Konzepte / Komplexe Beschreibungen
- Definition verfügbar
- notwendige und hinreichende Bedingungen
- basierend auf
  - (primitiven) Konzepten
  - Relationen / Rollen
  - Konzeptbildungsoperatoren
- implizite Einordnung in Subsumptionshierarchie

Beschreibungslogiken

Andere Namen

- Terminologische Logiken, KL-ONE-artige Sprachen
- Description logic, terminological logics, taxonomic logics, term subsumption systems, KL-ONE-like systems

Definition eines logischen Systems: Generelles Schema (s. LOS)

- eine formale Sprache (zur Repräsentation)
- Evaluations- / Interpretationsprinzipien
- semantische Kategorisierungen und Beziehungen
- Ableitungs-, Beweisverfahren

➢ **Beschreibungslogiken bilden logische Systeme**

Vokabular von Beschreibungslogiken

Typen von Ausdrücken

- Beschreibungen / Konzepte
- Rollen
- Aussagen/Formeln (Wissensbasis ist Menge von Formeln)

‚logische Symbole' (feststehende Interpretation)

- Konzeptbildungsoperatoren
- Rollenbildungsoperatoren
- Operatoren zur Bildung von Aussagen / Formeln

frei verfügbare Symbole

- Konzeptsymbole
- Rollensymbole
- (Konstanten)

Hilfssymbole

- Klammern, Interpunktion
- Symbole für natürliche Zahlen: 1, 2, 3, …

Beschreibungslogiken: Varianten

Unterschiede zwischen verschiedenen Beschreibungslogiken

- Auswahl von Operatoren zur Bildung von Aussagen
- Berücksichtigung von Konstanten (für Objekte)
- Auswahl von Konzeptbildungsoperatoren
- Auswahl von Rollenbildungsoperatoren

Auswirkungen

- Manchmal keine: Einschränkung von Formulierungsvarianten desselben Inhaltes
- Reine Konzeptsysteme vs. Konzeptsysteme + Weltausschnitt
- Ausdrucksmächtigkeit vs. Verarbeitbarkeit

Beschreibungslogik: $\mathcal{DL}$ (B & L, 2001): Symbole

(eine einfache Beschreibungslogik)

Freies Vokabular

- Atomare Konzepte: Lebewesen, Student, Mutter
- Rollen: hatKind, veranstalterIn
- Konstanten: snoopy, wbs, los

Festes Vokabular

- Vorgegebenes atomares Konzept: Thing
- Konzeptbildung: **ALL**, **EXISTS**, **FILLS**, **AND**
- Aussagen- / Formelbildung: $\sqsubseteq$, $\doteq$, $\rightarrow$

Hilfssymbole

- eckige und runde Klammern

Beschreibungslogik: $\mathcal{DL}$ (B & L, 2001): Grammatik

Konzepte (Beschreibungen)

- Jedes atomare Konzept ist ein Konzept.
- Wenn r eine Rolle, n eine nat. Zahl, und c eine Konstante ist sowie d, d_1, …, d_i Konzepte sind, dann sind
  - [**ALL** r d]
  - [**EXISTS** n r]
  - [**FILLS** r c]
  - [**AND** d_1, …, d_i]

(komplexe) Konzepte, und es gibt keine weiteren Konzepte.

Aussagen / Formeln

- Wenn c eine Konstante ist sowie d_1, d_2 Konzepte sind, dann sind $(d_1 \sqsubseteq d_2)$, $(d_1 \doteq d_2)$ und $(c \rightarrow d_1)$ Formeln (und das sind alle Formeln).

Bedeutung / Interpretation: Typenbeschränkungen

Konstanten

- werden durch individuelle Objekte interpretiert

Atomare und komplexe Konzepte

- werden durch Objektmengen (Kategorien) interpretiert

Rollen

- werden durch Mengen von Paaren (Relationen) interpretiert.

Aussagen / Formeln

- werden durch Wahrheitswerte interpretiert

$\mathcal{DL}$: Intendierte Bedeutungen fester Symbole

Thing: Alle Objekte / Dinge

Aussagen- / Formelbildung

$(d_1 \sqsubseteq d_2)$:	d_1 wird von d_2 subsumiert.
$(d_1 \doteq d_2)$:	d_1 und d_2 sind gleich.
$(c \rightarrow d_1)$:	c ist Instanz von d_1

Komplexe Konzepte

[**ALL** r d]:	Objekte, die nur zu Instanzen von d in Relation r stehen
[**EXISTS** n r]:	Objekte, die zu mindestens n Objekten in Relation r stehen
[**FILLS** r c]:	Objekte, die zu Objekt c in Relation r stehen
[**AND** d_1, …, d_i]:	gemeinsame Instanzen der Konzepte d_1, …, d_i

Beispiel: Konzeptsysteme (in $\mathcal{DL}$)

Wissensbasis

Lebewesen $\sqsubseteq$ Thing

Mensch $\sqsubseteq$ Lebewesen

weiblich $\sqsubseteq$ Lebewesen

Frau $\doteq$ [**AND** weiblich Mensch]

männlich $\sqsubseteq$ Lebewesen ; weitere Bedingungen in $\mathcal{DL}$ nicht ausdrückbar

Mann $\doteq$ [**AND** männlich Mensch]

Elternteil $\doteq$ [**AND** Mensch [**EXISTS** 1 hatKind]]

Mutter $\doteq$ [**AND** weiblich Elternteil]

Vater $\doteq$ [**AND** männlich Elternteil]

Großmutter $\sqsubseteq$ [**AND** Frau [**EXISTS** 1 hatKind]] ; weitere Bedingungen in $\mathcal{DL}$ nicht ausdrückbar

Aufgabe: Spezifizieren Sie ein Konzeptsystem

Domäne: Fahrzeugvermietung

- Fahrzeug, Bauteil
- Konzepte
  - Landfahrzeug, Wasserfahrzeug, Zweirad, Amphibienfahrzeug, Segelboot, Motorboot, LKW, PKW, Fahrrad, Motorrad, Benzinmotor, Dieselmotor, Motor, Segel, Ruder, Rad ...
  - DefektesFahrzeug, DefektesBauteil, Fahrbereit
- Rollen
  - hat-teil ...
- Konstanten
  - pkw18, rad23, motor18 ...

Vorgehen

- Schreiben Sie zunächst in natürlicher Sprache auf, was Sie ausdrücken wollen
- Versuchen Sie dann dies in den Formalismus zu gießen.
- Notieren Sie, welche Bedingungen Sie nicht ausdrücken können.

$\mathcal{DL}$: Bedeutung / Interpretation: Formal

Interpretationen sind Paare $\mathfrak{I} = <\mathcal{D}, \mathcal{I}>$

$\mathcal{D}$: nicht-leere Menge von Objekten

$\mathcal{I}$: Abbildung von atomaren Konzepten und Rollen auf Teilmengen von $\mathcal{D}$ bzw. von $\mathcal{D}^2$

Fortsetzung von $\mathcal{I}$ für komplexe Beschreibungen

$\mathcal{I}(\text{Thing}) = \mathcal{D}$

$\mathcal{I}([\textbf{ALL}\ r\ d]) = \{x \in \mathcal{D} \mid \forall y\ [(x, y) \in \mathcal{I}(r) \Rightarrow y \in \mathcal{I}(d)]\}$

$\mathcal{I}([\textbf{EXISTS}\ n\ r]) = \{x \in \mathcal{D} \mid \exists_n\ y\ [(x, y) \in \mathcal{I}(r)]\}$

$\mathcal{I}([\textbf{FILLS}\ r\ c]) = \{x \in \mathcal{D} \mid (x, \mathcal{I}(c)) \in \mathcal{I}(r)\}$

$\mathcal{I}([\textbf{AND}\ d_1, \ldots, d_i]) = \mathcal{I}(d_1) \cap \ldots \cap \mathcal{I}(d_i)$

$\mathcal{DL}$: Bedeutung / Interpretation: Formal (Forts.)

Interpretationen sind Paare $\mathfrak{I} = <\mathcal{D}, \mathcal{I}>$

Fortsetzung von $\mathcal{I}$ für Formeln

$\mathcal{I}((d_1 \sqsubseteq d_2))$ = wahr genau dann, wenn $\mathcal{I}(d_1) \subseteq \mathcal{I}(d_2)$.

$\mathcal{I}((d_1 \doteq d_2))$ = wahr genau dann wahr, wenn $\mathcal{I}(d_1) = \mathcal{I}(d_2)$.

$\mathcal{I}((c \rightarrow d_1))$ = wahr genau dann wahr, wenn $\mathcal{I}(c) \in \mathcal{I}(d_1)$.

Wir schreiben dann auch

$\mathfrak{I} \vDash (d_1 \sqsubseteq d_2)$ bzw. $\mathfrak{I} \vDash (d_1 \doteq d_2)$ bzw. $\mathfrak{I} \vDash (c \rightarrow d_1)$

und sagen

$\mathfrak{I}$ ist ein Modell der Aussage / Formel.

$\mathfrak{I}$ macht die Aussage / Formel wahr.

Äquivalenz und Subsumption

Definitionen

Es seien d_1 und d_2 Beschreibungen.

- d_1 subsumiert genau dann d_2, wenn für jede Interpretation $\mathfrak{I} = <\mathcal{D}, \mathcal{I}>$ gilt: $\mathcal{I}(d_1) \subseteq \mathcal{I}(d_2)$.
- d_1 und d_2 sind genau dann äquivalent, wenn für jede Interpretation $\mathfrak{I} = <\mathcal{D}, \mathcal{I}>$ gilt: $\mathcal{I}(d_1) = \mathcal{I}(d_2)$.

Korollar

- d_1 subsumiert genau dann d_2, wenn für jede Interpretation $\mathfrak{I}$ gilt: $\mathfrak{I} \vDash (d_1 \sqsubseteq d_2)$
- d_1 und d_2 sind genau dann äquivalent, wenn für jede Interpretation $\mathfrak{I}$ gilt: $\mathfrak{I} \vDash (d_1 \doteq d_2)$.

In entsprechenden Fällen schreiben wir auch

- $\vDash (d_1 \sqsubseteq d_2)$ bzw. $\vDash (d_1 \doteq d_2)$. (Die Formeln sind gültig.)

Beispiel: Subsumption und Äquivalenz von Beschreibungen

$\models$ ([**AND** weiblich Elternteil] $\sqsubseteq$ Elternteil)

$\models$ (Elternteil $\sqsubseteq$ Thing)

$\models$ ([**AND** weiblich Elternteil] $\doteq$ [**AND** Elternteil weiblich])

$\models$ ([**AND** Thing Elternteil] $\doteq$ Elternteil)

$\models$ ([**ALL** hatKind [AND weiblich Elternteil]] $\sqsubseteq$ [**ALL** hatKind weiblich])

$\models$ ([**ALL** hatKind Thing] $\doteq$ Thing)

$\models$ ([**EXISTS** 3 hatKind] $\sqsubseteq$ [**EXISTS** 2 hatKind])

$\models$ ([**FILLS** hatKind harry] $\sqsubseteq$ [**EXISTS** 1 hatKind])

Erfüllbarkeit und Folgerung

Definitionen

Es sei X eine Formel und KB eine Menge von Formeln.

- Eine Interpretation $\Im$ erfüllt genau dann KB, wenn sie alle Elemente von KB wahr macht.
- KB ist genau dann erfüllbar, wenn es eine Interpretation $\Im$ gibt, die sie erfüllt.
- X folgt genau dann aus KB, wenn jede Interpretation $\Im$, die KB erfüllt, auch X wahr macht.

Wir schreiben für Folgerung auch

- In Zeichen: KB $\models$ X

Beispiel: Folgerung aus einer Wissensbasis

KB = { Mensch $\sqsubseteq$ Lebewesen, weiblich $\sqsubseteq$ Lebewesen,
Frau $\doteq$ [**AND** weiblich Mensch],
Elternteil $\doteq$ [**AND** Mensch [**EXISTS** 1 hatKind]],
Mutter $\doteq$ [**AND** weiblich Elternteil],
Großmutter $\sqsubseteq$ [**AND** Frau [**EXISTS** 1 hatKind]] }

KB $\models$ (Mutter $\sqsubseteq$ Lebewesen)

KB $\models$ (Großmutter $\sqsubseteq$ Mutter)

$\mathcal{DL}$: Berechnung von Folgerung: Struktureller Ansatz

Grundannahmen

  - sind nicht aus semantischen Gründen erforderlich
  - für effiziente Verarbeitung wichtig
  - beeinträchtigen die Ausdrucksmächtigkeit
- Auf der linken Seite von ($d_1 \sqsubseteq d_2$) bzw. ($d_1 \doteq d_2$) Formeln in der Wissensbasis stehen nur (freie) atomare Konzepte.
- Kein atomares Konzept taucht mehr als einmal auf der linken Seite so einer Formel auf.
- Die Wissensbasis enthält keinen Zyklus.
  - Die Formeln lassen sich so ordnen, dass kein atomares Konzept auf einer linken Seite auftritt, nachdem es schon auf einer rechten Seite verwendet wurde.

Eine weitere mögliche Grundannahme

Häufig (aber nicht hier) vorausgesetzt

- Die Wissensbasis enthält keine Formeln der Form $(d_1 \sqsubseteq d_2)$
- Ersatzmöglichkeit:
  $(d_1 \doteq [\textbf{AND}\ d_2\ d_1^*])$
  wobei d_1^* ein ansonsten ungenutztes atomares Konzept ist.

Berechnung von Folgerung: Struktureller Ansatz

Aufgabe

gegeben KB (wie eben eingeschränkt) und zwei beliebige (komplexe) Beschreibungen d_1 und d_2
bestimme, ob KB $\models (d_1 \sqsubseteq d_2)$.

Vorgehen

1) Äquivalenzaussagen werden genutzt, um normalisierte Formen von d_1 und d_2 zu bilden (d_1' und d_2').
   - Es gilt dann: KB $\models (d_1 \doteq d_1')$ und KB $\models (d_2 \doteq d_2')$
2) Subsumptionsaussagen und Strukturähnlichkeit werden genutzt, um KB $\models (d_1' \sqsubseteq d_2')$ zu bestimmen
   - Semantik sichert dann dasselbe Ergebnis für KB $\models (d_1 \sqsubseteq d_2)$ zu

$\mathcal{DL}$: Normalisierung einer Beschreibung d

1) **Expandiere Definitionen**
   - Ersetze in d jedes atomare Konzept a, für das in KB eine Aussage $a \doteq e$ auftritt, durch e
2) **Verflache AND-Schachtelungen**
   - Ersetze [**AND** … [**AND** d_1 … d_i] …] durch [**AND** … d_1 … d_i …]
3) **Kombiniere ALL-Operatoren**
   - Ersetze [**AND** … [**ALL** r d_1] … [**ALL** r d_2] …] durch [**AND** … [**ALL** r [**AND** d_1 d_2]] … …]
4) **Kombiniere EXISTS-Operatoren**
   - Ersetze [**AND** … [**EXISTS** n_1 r] … [**EXISTS** n_2 r] …] mit $n = \max(n_1, n_2)$ durch [**AND** … [**EXISTS** n r] … …]

$\mathcal{DL}$: Normalisierung einer Beschreibung d

5) **Reduziere Thing**
   - Lösche Thing aus jeder Konjunktion.
   - Ersetze die leere Konjunktion [**AND**] durch Thing
   - Ersetze die leere Einschränkung [**ALL** r Thing] durch Thing
6) **Lösche redundante Ausdrücke**
   - Lösche mehrfach-Kopien eines Ausdrucks in Konjunktionen.
7) **Normalisiere atomare Konzepte**
   - Wenn schließlich nur ein atomares Konzept $d \neq$ Thing bleibt, gib [**AND** d] zurück.

$\mathcal{DL}$: Resultat der Normalisierung einer Beschreibung

Thing oder eine Beschreibung der Form

[AND $a_1 \dots a_m$
[FILLS r_1 c_1] … **[FILLS** $r_{m'}$ $c_{m'}$]
[EXISTS n_1 s_1] … **[EXISTS** $n_{m''}$ $s_{m''}$]
[ALL t_1 e_1] … **[ALL** $t_{m'''}$ $e_{m'''}$]]

wobei

$a_1 \dots a_m$ verschiedene, atomare, nicht-definierte Konzepte sind

$s_1 \dots s_{m''}$ verschiedene Rollen sind und

$t_1 \dots t_{m'''}$ verschiedene Rollen sind

$\mathcal{DL}$: Beispiel: Normalisierung

KB = { Mensch ⊑ Lebewesen, weiblich ⊑ Lebewesen,
Frau ≐ **[AND** weiblich Mensch],
Elternteil ≐ **[AND** Mensch **[EXISTS** 1 hatKind]],
Mutter ≐ **[AND** weiblich Elternteil],
Großmutter ⊑ **[AND** Frau **[EXISTS** 1 hatKind]] }

- Mutter ⇝ **[AND** weiblich Elternteil]
  ⇝ **[AND** weiblich **[AND** Mensch **[EXISTS** 1 hatKind]]]
  ⇝ **[AND** weiblich Mensch **[EXISTS** 1 hatKind]]
- Großmutter ⇝ **[AND** Großmutter]
- **[AND** Frau **[EXISTS** 1 hatKind]]
  ⇝ **[AND [AND** weiblich Mensch] **[EXISTS** 1 hatKind]]
  ⇝ **[AND** weiblich Mensch **[EXISTS** 1 hatKind]]

$\mathcal{DL}$: Beispiel: Normalisierung

KB = {Techie ≐ **[EXISTS** 2 techDegree],
WellRoundedCo
≐ **[AND** Company
[ALL manager **[AND** B-SchoolGrade
[EXISTS 1 techDegree]]]],
HighTechCo
≐ **[AND** Company
[FILLS exchange Nasdaq]
[ALL manager Techie]] }

[AND WellRoundedCo HighTechCo]

⇝ **[AND [AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 1 techDegree]]]] **[AND** Company **[FILLS** exchange Nasdaq] **[ALL** manager **[EXISTS** 2 techDegree]]]]

$\mathcal{DL}$: Beispiel: Normalisierung

[AND WellRoundedCo HighTechCo]

⇝ **[AND [AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 1 techDegree]]]] **[AND** Company **[FILLS** exchange Nasdaq] **[ALL** manager **[EXISTS** 2 techDegree]]]]

⇝ **[AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 1 techDegree]]] Company **[FILLS** exchange Nasdaq] **[ALL** manager **[EXISTS** 2 techDegree]]]

⇝ **[AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 1 techDegree] **[EXISTS** 2 techDegree]]] Company **[FILLS** exchange Nasdaq]]

⇝ **[AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 2 techDegree]]] Company **[FILLS** exchange Nasdaq]]

⇝ **[AND** Company **[ALL** manager **[AND** B-SchoolGrade **[EXISTS** 2 techDegree]]] **[FILLS** exchange Nasdaq]]

Strukturvergleich – Structure Matching

Grundidee

- KB $\models (d_1 \sqsubseteq d_2)$ gilt genau dann, wenn d_2 = Thing oder es zu jeder Bedingung in der Konjunktion d_2 eine Bedingung in d_1 gibt, die gleich oder stärker einschränkt.
- Die Normalisierung erleichtert das Finden korrespondierender Bedingungen

Achtung:

- Strukturvergleich funktioniert für die hier vorgestellte Sprache $\mathcal{DL}$. Bei anderen Beschreibungslogiken können Interaktionen zwischen Bedingungen weitere Einschränkungen hervorrufen.

$\mathcal{DL}$: Strukturvergleich: Vorgehen

Eingabe

- zwei normalisierte Beschreibungen d = [**AND** $d_1 \ldots d_m$] und e = [**AND** $e_1 \ldots e_{m'}$] und eine Wissensbasis KB.

Aufgabe: Bestimmung, ob KB $\models (d \sqsubseteq e)$

Für i := 1 bis m'

Wenn es keine e_i bzgl. KB spezialisierende Bedingung d_j gibt, dann bricht ab, e subsumiert d nicht

Ende (Für i)

e subsumiert d

Ausdrucksstärke und Berechnungsaufwand

Beschreibungslogiken

- unterscheiden sich in der Menge der Konzeptbildungsoperatoren
- sind ein Beispiel für eine Sprachfamilie, für die die Abhängigkeit zwischen Ausdrucksstärke und Berechnungsaufwand sehr detailliert untersucht ist

Nächste Sitzung

- Mehr technische Details
- Mehr Konzeptbildungsoperatoren
- Überblick über die Sprachfamilie