Wissensrepräsentation

–

Christopher Habel, Hedda Schmidtke

Sommersemester 2004

Sitzung 8: Beschreibungslogiken (2)

- Fortsetzung Vorlesung 7
- Übersicht über die Sprachfamilie
- Terminologien (T-Box) und Weltmodell (A-Box)

$\mathcal{DL}$: Strukturvergleich: Vorgehen

Eingabe

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

Aufgabe: Bestimmung, ob KB ⊨ (d ⊑ 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

$\mathcal{DL}$: Strukturvergleich: spezialisierende Bedingungen

d_i spezialisiert Bedingung e_j bzgl. KB falls

- d_i = e_j ist

oder

- e_j = [**EXISTS** n r] und d_i = [**EXISTS** n' r] mit n' > n
- e_j = [**EXISTS** 1 r] und d_i = [**FILLS** r c]
- e_j = [**ALL** r e'] und d_i = [**ALL** r d'] und KB ⊨ (d' ⊑ e')
  - einfacher zu prüfen, da d' und e' schon normalisiert sind und einfacher als d_i und e_j.

oder

- d_i atomar ist, (d_i ⊑ d') ∈ KB und KB ⊨ (d' ⊑ e_j)
  - erfordert Normalisierung von d'

$\mathcal{DL}$: Beispiel: Strukturvergleich

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

Aufgabe: KB ⊨ (Mutter ⊑ Lebewesen)

- Mutter ⇝ [**AND** weiblich Mensch [**EXISTS** 1 hatKind]]
- Lebewesen ⇝ [**AND** Lebewesen]
- weiblich und Mensch spezialisieren beide Lebewesen

Aufgabe: KB ⊨ (Großmutter ⊑ Mutter)

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

$\mathcal{DL}$: Klassifikation und Äquivalenzberechnung

Klassifikation

- Bestimme, ob KB $\models$ (c → e)
- Es sei d = [**AND** d_1 … d_m], wobei die d_i genau die Beschreibungen sind, für die (c → d_i) ∈ KB. Dann gilt:
- KB $\models$ (c → e) gdw. KB $\models$ (d ⊑ e)

Äquivalenz

- Bestimme ob KB $\models$ (d ≐ e)
- KB $\models$ (d ≐ e) gdw. KB $\models$ (d ⊑ e) und KB $\models$ (e ⊑ d)

Der Strukturvergleich

- kann auch zur Berechnung von Klassifikation und Äquivalenz eingesetzt werden

$\mathcal{DL}$: Korrektheit (und Vollständigkeit) der Strukturbasierten Berechnung

Normalisierung

- Jeder Schritt ist eine Äquivalenzumformung.
  - Ersetzungstheorem gilt.
  - Basisäquivalenzen ergeben sich aus der Festlegung der Interpretation.

Strukturvergleich

- Wenn eine Bedingung d eine Bedingung e bzgl. einer Wissensbasis spezialisiert, dann gilt auch KB $\models$ (d ⊑ e)
  - Ergibt sich für die Einzelbedingungen (über strukturelle Induktion) aus der Festlegung der Interpretation.
- Für $\mathcal{DL}$ gilt: Wenn es eine Teil-Bedingung von e gibt, zu der es keine sie bzgl. KB spezialisierende Teil-Bedingung von d gibt, dann gilt KB $\models$ (d ⊑ e) nicht.➜ [B & L, 2001]

Grenzen des Strukturvergleichs

Spracherweiterung

- [**AT-MOST** n r]: Intendierte Bedeutung: trifft auf x zu, wenn x höchstens zu n Objekten in der Relation r steht.

Beispiel

- [**ALL** r d] subsummiert
  [**AND** [**FILLS** r c] [**AT-MOST** 1 r] [**ALL** s d] [**FILLS** s c]]
- dies ist aber durch einen einfachen Strukturvergleich nicht ersichtlich.

Literatur

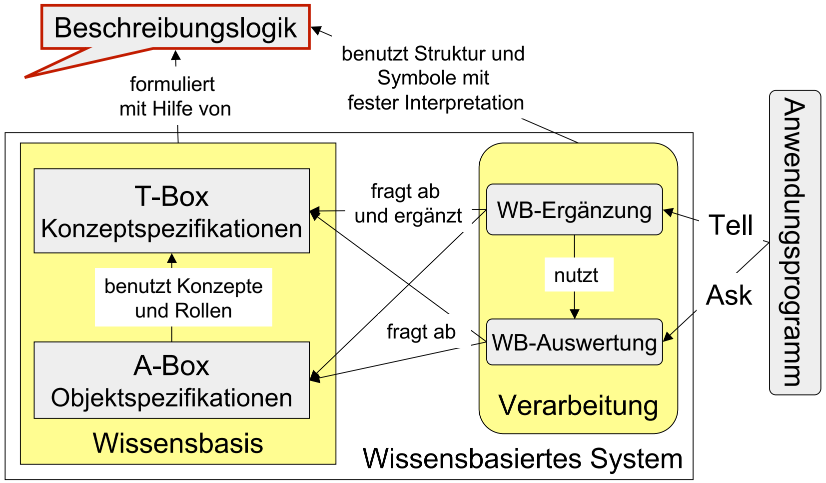
F. Baader, D. Calvansee, D.L. McGuinness, D. Nardi & P. Patel-Schneider (eds.) The Description Logic Handbook. Theory, Implementation and Application. Cambridge UP: Cambridge, NY.

Insbesondere

Baader, Franz & Werner Nutt (2003). Kapitel 2. Basic description logics.

Baader, Franz (2003). Appendix. Description logic terminology.

Wissensbasiertes System mit Beschreibungslogiken



Mögliche Anwendungen von Beschreibungslogiken

Aufbau von Konzeptsystemen

- inkrementelle Anreicherung
- Wissenserwerb
  - Aufdeckung impliziter Subsumptionsbeziehungen
  - insb. nicht-intendierte

Aufbau von Weltmodellen

- Konfiguration
  - Ausschluss von Designs, die Konflikte enthalten
- Diagnose und Überwachung
  - Erkennung von Problemen auf Basis ihrer Beschreibungen
- Ausdrucksreichere Produktionssysteme
- Datenbankanfragen

Konzeptbildungsoperatoren (Concept constructors)

Standardnotation der Konstruktoren

$\top$: das universelle Konzept (Thing)
$C \sqcap D$: Durchschnitt, Bedingungskonjunktion ([**AND** C D])
$C \sqcup D$: Vereinigung, Bedingungsdisjunktion
$\neg A$: Komplement, Negation nur atomarer Konzepte
$\neg C$: Komplement, Negation nur beliebige Konzepte
$\bot$: das leere Konzept
$\forall R.C$: Werterestriktion ([**ALL** r d])
$\exists R.\top$: eingeschränkte Existenzrestriktion ([**EXISTS** 1 r])
$\exists R.D$: freie Existenzrestriktion
$(\geq n\ R)$: Anzahlrestriktion mindestens ([**EXISTS** n r])
$(\leq n\ R)$: Anzahlrestriktion höchstens

Interpretation der Konstruktoren

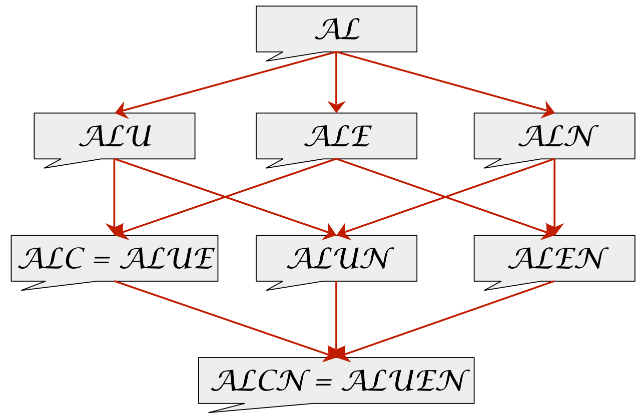
Interpretationen sind Paare $\mathfrak{I} = \langle \mathcal{D}, \mathcal{I} \rangle$

$\mathcal{I}(\top) = \mathcal{D}$
$\mathcal{I}(C \sqcap D) = \mathcal{I}(C) \cap \mathcal{I}(D)$
$\mathcal{I}(C \sqcup D) = \mathcal{I}(C) \cup \mathcal{I}(D)$
$\mathcal{I}(\neg C) = \mathcal{D} \setminus \mathcal{I}(C)$
$\mathcal{I}(\bot) = \emptyset$
$\mathcal{I}(\forall R.C) = \{a \in \mathcal{D} \mid \forall b\ [(a, b) \in \mathcal{I}(R) \Rightarrow b \in \mathcal{I}(C)]\}$
$\mathcal{I}(\exists R.\top) = \{a \in \mathcal{D} \mid \exists b\ [(a, b) \in \mathcal{I}(R)]\}$
$\mathcal{I}(\exists R.D) = \{a \in \mathcal{D} \mid \exists b\ [(a, b) \in \mathcal{I}(R) \wedge b \in \mathcal{I}(D)]\}$
$\mathcal{I}((\geq n\ R)) = \{a \in \mathcal{D} \mid\ |\{b \in \mathcal{D} \mid (a, b) \in \mathcal{I}(R)\}| \geq n\}$
$\mathcal{I}((\leq n\ R)) = \{a \in \mathcal{D} \mid\ |\{b \in \mathcal{D} \mid (a, b) \in \mathcal{I}(R)\}| \leq n\}$

Die Basissprache $\mathcal{AL}$ und Erweiterungen

	$\mathcal{DL}$	$\mathcal{AL}$	$\mathcal{FL}^-$	$\mathcal{FL}_0$	$\mathcal{ALU}$	$\mathcal{ALE}$	$\mathcal{ALN}$	$\mathcal{ALC}$
$\top$	+	+	+	+	+	+	+	+
$C \sqcap D$	+	+	+	+	+	+	+	+
$C \sqcup D$	–	–	–	–	+	–	–	–
$\neg A$	–	+	–	–	+	+	+	+
$\neg C$	–	–	–	–	–	–	–	+
$\bot$	–	+	–	–	+	+	+	+
$\forall R.C$	+	+	+	+	+	+	+	+
$\exists R.\top$	+	+	+	–	+	+	+	+
$\exists R.D$	–	–	–	–	–	+	–	–
$(\geq n\ R)$	+	–	–	–	–	–	+	–
$(\leq n\ R)$	–	–	–	–	–	–	+	–

$\mathcal{AL}$-Sprachfamilie (Ausdrucksmächtigkeit)



Rollenkonstruktoren mit Interpretation

$\mathcal{I}(U) = \quad \mathcal{D} \times \mathcal{D}$

$\mathcal{I}(Id) = \quad \{(a, a) \mid a \in \mathcal{D}\}$

$\mathcal{I}(R \sqcap S) = \quad \mathcal{I}(R) \cap \mathcal{I}(S)$

$\mathcal{I}(R \sqcup S) = \quad \mathcal{I}(R) \cup \mathcal{I}(S)$

$\mathcal{I}(\neg R) = \quad \mathcal{D} \times \mathcal{D} \setminus \mathcal{I}(R)$

$\mathcal{I}(R^-) = \quad \{(a, b) \in \mathcal{D} \times \mathcal{D} \mid (b, a) \in \mathcal{I}(R)\}$

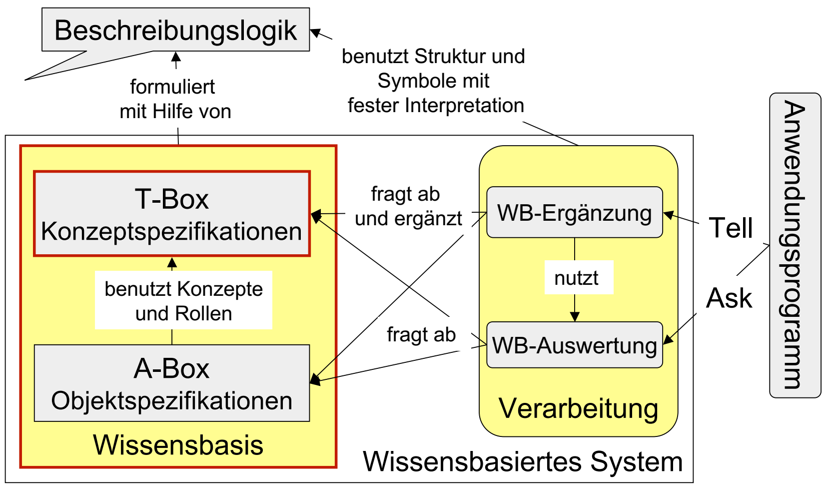
$\mathcal{I}(R \circ S) = \quad \{(a, b) \in \mathcal{D} \times \mathcal{D} \mid \exists c\ [(a, c) \in \mathcal{I}(R) \wedge (c, b) \in \mathcal{I}(S)]\}$

$\mathcal{I}(R^+) =$ Transitive Hülle von $\mathcal{I}(R)$

$\mathcal{I}(R^*) =$ Reflexive, transitive Hülle von $\mathcal{I}(R)$

$\mathcal{I}(R|_C) = \quad \mathcal{I}(R) \cap \mathcal{D} \times \mathcal{I}(C)$

Wissensbasiertes System mit Beschreibungslogiken



Struktur der Wissensbasis

T-Box (Terminologie)

- enthält Formeln der Form $(C \sqsubseteq D)$, $(C \doteq D)$, wobei C und D Konzepte sind.
- ggf. auch Formeln der Form $(R \sqsubseteq S)$, $(R \doteq S)$, wobei R und S Rollen(beschreibungen) sind.

A-Box (Assertionen, Weltmodell)

- enthält Formeln der Form C(a) (bei L&B, 2001: $(a \rightarrow C)$) und R(a, b), wobei a und b Konstanten, C ein Konzept und R eine Rolle ist.

Definitorische Terminologien

T-Box

- erlaubt die Zuordnung von Namen zu komplexen Beschreibungen (Definitionen: $(A \doteq D)$)
- T-Box-Bedingung: kein Name ist mehrfach definiert

Vokabular

- Konzepte und Rollen
- entsprechend der Definitionen in der T-Box lässt sich unterscheiden:
  - Basissymbole
  - definierte Symbole

Interpretation definitorischer Terminologien

Basisinterpretationen

- weisen nur den Basis-Symbolen Werte zu

Eine definitorische T-Box

- ist eine T-Box, die für jede Basisinterpretation genau eine Erweiterung zulässt, die Modell der T-Box ist.
- Die T-Box und die Basisinterpretationen bestimmen alle Interpretationen.

Beobachtung: Erweiterung einer Basisinterpretation

- Eine Definition, die nur Basissymbole verwendet, erlaubt genau eine Erweiterung der Interpretation, die die Definition erfüllt.
- Entsprechendes gilt allgemeiner für azyklische Definitionen

Ergebnisse

- Jede azyklische T-Box ist definitorisch.
- Zu jeder definitorischen T-Box gibt es eine äquivalente azyklische T-Box.

Zyklische Terminologien: Beispiel

Beispiel

- Ein Mensch ist ein Lebewesen mit menschlichen Eltern.

$\text{Human} \doteq \text{Animal} \sqcap \forall \text{hasParent.Human}$

- Ein Mann nur mit männlichen Nachkommen

$\text{Momo} \doteq \text{Man} \sqcap \forall \text{hasChild.Momo}$

- Eine Wurzel eines binären Baumes ist eine Baumwurzel, hat höchstens zwei Äste und alle Äste sind binäre Bäume.

$\text{binTreeR} \doteq \text{TreeR} \sqcap {\leq}2\ \text{hasBranch} \sqcap \forall \text{hasBranch.binTreeR}$

Zyklische Definition: Beispiel (1)

T-Box = {Momo $\doteq$ Man $\sqcap$ $\forall$hasChild.Momo}

$\mathcal{I}$(Man) = $\mathcal{D}$ = {$Hans_1$, $Hans_2$, …, $Hans_5$,$Karl_1$, $Karl_2$, …}

$\mathcal{I}$(hasChild)={($Hans_i$, $Hans_{i+1}$) | $1 \leq i \leq 4$} $\cup$ {($Karl_i$, $Karl_{i+1}$) | $1 \leq i$}

Beobachtung zu Man $\sqcap$ $\forall$hasChild.Momo

$Hans_5 \in \mathcal{I}$(Man $\sqcap$ $\forall$hasChild.Momo), unabhängig von der (noch nicht feststehenden) Interpretation von Momo.

Wenn die T-Box erfüllt werden soll, muss also auch gelten: $Hans_5 \in \mathcal{I}$(Momo)

Über die anderen wissen wir erstmal noch nichts.

Erweiterung von $\mathcal{I}$

$\mathcal{I}_1$(Man) = $\mathcal{I}$(Man), $\mathcal{I}_1$(hasChild) = $\mathcal{I}$(hasChild)

$\mathcal{I}_1$(Momo) = {$Hans_5$}

Zyklische Definition: Beispiel (2)

T-Box = {Momo $\doteq$ Man $\sqcap$ $\forall$hasChild.Momo}

$\mathcal{I}_1$(Man) = $\mathcal{D}$ = {$Hans_1$, $Hans_2$, …, $Hans_5$,$Karl_1$, $Karl_2$, …}

$\mathcal{I}_1$(hasChild)={($Hans_i$, $Hans_{i+1}$) | $1 \leq i \leq 4$} $\cup$ {($Karl_i$, $Karl_{i+1}$) | $1 \leq i$}

$\mathcal{I}_1$(Momo) = {$Hans_5$}

Beobachtung zu Man $\sqcap$ $\forall$hasChild.Momo

$Hans_4 \in \mathcal{I}_1$(Man $\sqcap$ $\forall$hasChild.Momo), unabhängig von der (noch nicht feststehenden) Interpretation von Momo.

Wenn die T-Box erfüllt werden soll, muss also auch gelten: $Hans_4 \in \mathcal{I}_1$(Momo)

Erweiterung von $\mathcal{I}_1$

$\mathcal{I}_2$(Man) = $\mathcal{I}$(Man), $\mathcal{I}_2$(hasChild) = $\mathcal{I}$(hasChild)

$\mathcal{I}_2$(Momo) = $\mathcal{I}_1$(Momo) $\cup$ {$Hans_4$}

Zyklische Definition: Beispiel (3)

Entsprechende Überlegungen führen dann zu

T-Box = {Momo $\doteq$ Man $\sqcap$ $\forall$hasChild.Momo}

$\mathcal{I}_5$(Man) = $\mathcal{D}$ = {$Hans_1$, $Hans_2$, …, $Hans_5$,$Karl_1$, $Karl_2$, …}

$\mathcal{I}_5$(hasChild)={($Hans_i$, $Hans_{i+1}$) | $1 \leq i \leq 4$} $\cup$ {($Karl_i$, $Karl_{i+1}$) | $1 \leq i$}

$\mathcal{I}_5$(Momo) = {$Hans_1$, $Hans_2$, …, $Hans_5$}

Beobachtung zu Man $\sqcap$ $\forall$hasChild.Momo

$\mathcal{I}_5$(Momo) = $\mathcal{I}_5$(Man $\sqcap$ $\forall$hasChild.Momo)

Eine weitere Iteration ist nicht erforderlich.

Fixpunktsemantik

beruht auf der Definition solcher stabilen Modelle (als Fixpunkte von Funktionen)

Zyklische Terminologien: Fixpunktsemantik

Fixpunkt einer Basisinterpretation

- muss nicht eindeutig sein
- Beispiel: Es gibt zwei Erweiterungen von $\mathcal{I}$, die die Formel wahr machen.

  $\mathcal{I}'$(Momo) = {$Hans_1$, $Hans_2$, …, $Hans_5$}

  $\mathcal{I}''$(Momo) = $\mathcal{D}$

Ordnung

- Fixpunkt-Interpretationen ausgehend von derselben Basisinterpretation lassen sich partiell ordnen.
- $\mathcal{I}' \preceq \mathcal{I}''$, denn $\mathcal{I}'$(Momo) $\subseteq$ $\mathcal{I}''$(Momo)

➢ Least-/greatest fixpoint semantics

Zyklische Terminologien: Existenz von Fixpunkten

Beispiel: Kein Fixpunkt

T-Box = $\{A \doteq \neg A\}$

(da kein Modell)

Beispiel: ungeordnet Fixpunkte

T-Box = $\{A \doteq \forall R.\neg A\}$

Basisinterpretation: $\mathcal{D} = \{a, b\}$, $\mathcal{I}(R)=\{(a, b), (b, a)\}$

Fixpunkte: $\mathcal{I}'(A)=\{a\}$ $\mathcal{I}''(A)=\{b\}$

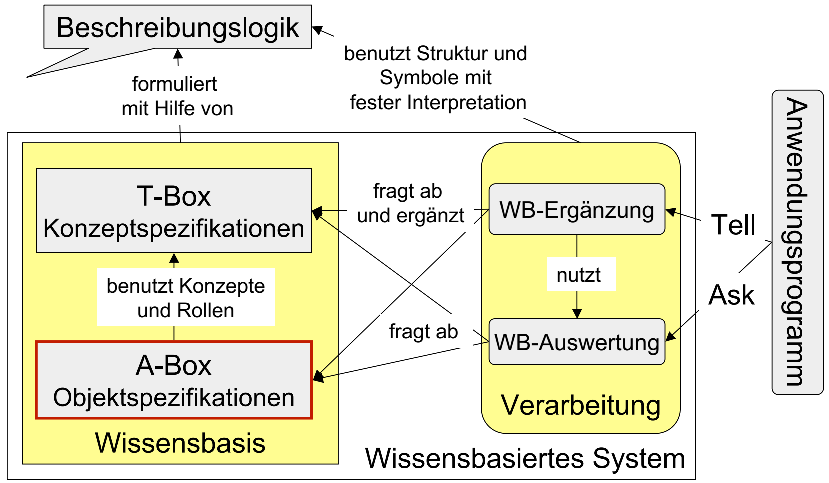
es gibt keinen kleinsten oder größten Fixpunkt bei dieser Basisinterpretation

Zyklische Terminologien

mit kleinsten und größten Fixpunkten

- lassen sich syntaktisch charakterisieren
- (über Anzahl der Negationen in einem Zyklus)

Wissensbasiertes System mit Beschreibungslogiken



A-Box und Schnittstellen zur T-Box

A-Box (Assertionen, Weltmodell)

- enthält Formeln der Form C(a) (bei L&B, 2001: $(a \rightarrow C)$) und R(a, b), wobei a und b Konstanten, C ein Konzept und R eine Rolle ist.

Interpretation < $\mathcal{I}$, $\mathcal{D}$ >

- jede Konstante wird durch $\mathcal{I}$ auf ein Objekt in $\mathcal{D}$ abgebildet.

$\mathcal{I}(C(a))$ = wahr, gdw. $\mathcal{I}(a) \in \mathcal{I}(C)$

$\mathcal{I}(R(a, b))$ = wahr, gdw. $(\mathcal{I}(a), \mathcal{I}(b)) \in \mathcal{I}(R)$

Konzeptbildungsoperatoren mit Zugriff auf A-Box (?)

'one of' : $\mathcal{I}(\{a_1, a_2, \ldots, a_n\}) = \{\mathcal{I}(a_1), \mathcal{I}(a_2), \ldots, \mathcal{I}(a_n)\}$

'fills': $\mathcal{I}(R{:}a) = \{b \in \mathcal{D} \mid (b, \mathcal{I}(a)) \in \mathcal{I}(R)\}$