

Beschreibungslogik

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

Sommersemester 2018

Thomas Schneider

AG Theorie der künstlichen Intelligenz (TdKI)

<http://tinyurl.com/ss18-bl>

Vorlesungsübersicht

Kapitel 1: Einleitung

Kapitel 2: Grundlagen

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

Kapitel 4: Tableau-Algorithmen

Kapitel 5: Komplexität

Kapitel 6: Effiziente Beschreibungslogiken

Kapitel 7: ABoxen und Anfragebeantwortung

Ziel des Kapitels

Die wichtigsten Eigenschaften einer BL sind

- **Ausdrucksstärke** und
- **Komplexität**

Die Ausdrucksstärke kann man **nicht linear quantifizieren**, sondern nur **beschreiben und charakterisieren**.

Wir werden nun

- mehrere Konstruktionen auf Modellen etablieren und
- diese verwenden, um die Ausdrucksstärke von BLen zu studieren.

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

- 1 Bisimulation
- 2 Endliche Modelleigenschaft und Filtration

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

1 Bisimulation

2 Endliche Modelleigenschaft und Filtration

Überblick

Bisimulation ist graphentheoretischer Begriff;
beschreibt die „Ähnlichkeit“ von Graphen.

Hängt eng zusammen mit der Ausdrucksstärke von \mathcal{ALC} .

Wir formalisieren den Zusammenhang und betrachten zwei
Anwendungen:

- Beweis der **Nicht-Ausdrückbarkeit** von Eigenschaften
- **Baummodelleigenschaft**

Bisimulation: Definition

Definition 3.1 (Bisimulation)

Seien \mathcal{I}_1 und \mathcal{I}_2 Interpretationen.

Relation $\rho \subseteq \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2}$ ist **Bisimulation** zwischen \mathcal{I}_1 und \mathcal{I}_2 , wenn gilt:

- 1 Wenn $d_1 \rho d_2$, dann gilt für alle Konzeptnamen A :

$$d_1 \in A^{\mathcal{I}_1} \quad \text{gdw.} \quad d_2 \in A^{\mathcal{I}_2}$$

- 2 Wenn $d_1 \rho d_2$ und $(d_1, d'_1) \in r^{\mathcal{I}_1}$ für beliebigen Rollennamen r , dann gibt es ein $d'_2 \in \Delta^{\mathcal{I}_2}$ mit $d'_1 \rho d'_2$ und $(d_2, d'_2) \in r^{\mathcal{I}_2}$.
- 3 Wenn $d_1 \rho d_2$ und $(d_2, d'_2) \in r^{\mathcal{I}_2}$ für beliebigen Rollennamen r , dann gibt es ein $d'_1 \in \Delta^{\mathcal{I}_1}$ mit $d'_1 \rho d'_2$ und $(d_1, d'_1) \in r^{\mathcal{I}_1}$.

T 3.1

Bisimulationstheorem

Seien \mathcal{I}_1 und \mathcal{I}_2 Interpretationen, $d_1 \in \Delta^{\mathcal{I}_1}$, $d_2 \in \Delta^{\mathcal{I}_2}$.

$(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$: es gibt Bisimulation ρ zwischen \mathcal{I}_1 und \mathcal{I}_2
mit $d_1 \rho d_2$ (wir sagen: d_1 ist **bisimilar** zu d_2).

Beachte: die leere Relation ist immer Bisimulation!

Theorem 3.2

Seien $\mathcal{I}_1, \mathcal{I}_2$ Interpretationen, $d_1 \in \Delta^{\mathcal{I}_1}$ und $d_2 \in \Delta^{\mathcal{I}_2}$.

Wenn $(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$, dann gilt für alle \mathcal{ALC} -Konzepte C :

$$d_1 \in C^{\mathcal{I}_1} \quad \text{gdw.} \quad d_2 \in C^{\mathcal{I}_2}$$

T 3.2

Intuitiv: \mathcal{ALC} kann nicht zwischen d_1 und d_2 „unterscheiden“.

Ausdrucksstärke

Wir interessieren uns für Eigenschaften von Elementen d
in Interpretationen \mathcal{I} .

Definition 3.3 (Eigenschaft, Ausdrückbarkeit)

Eine **Eigenschaft** E ist eine Menge von Paaren (\mathcal{I}, d) ,
wobei \mathcal{I} eine Interpretation und $d \in \Delta^{\mathcal{I}}$ ein Element in \mathcal{I} ist.

E ist **in \mathcal{ALC} ausdrückbar**, wenn es ein \mathcal{ALC} -Konzept C gibt,
so dass für alle \mathcal{I} und $d \in \Delta^{\mathcal{I}}$ gilt:

$$(\mathcal{I}, d) \in E \quad \text{gdw.} \quad d \in C^{\mathcal{I}}$$

Intuitionen:

- $d \in \Delta^{\mathcal{I}}$ hat die Eigenschaft E gdw. $(\mathcal{I}, d) \in E$.
- E ist ausdrückbar, wenn es ein **äquivalentes** \mathcal{ALC} -Konzept C gibt.

Ausdrucksstärke

Typische Eigenschaften,

für deren Ausdrückbarkeit wir uns interessieren:

- d hat einen r -Nachfolger in \mathcal{I}
- es gibt einen von d ausgehenden endlichen r -Pfad in \mathcal{I} , dessen letztes Element Instanz von C ist
- jedes Element von $\Delta^{\mathcal{I}}$ ist r -Nachfolger von d

Eigenschaft kann auch in logischem Formalismus gegeben sein, z. B.:

- $\exists r^- .A$
- $\exists y (r(y, x) \wedge A(y))$

Ausdrucksstärke und Bisimulation

Beschränkungen der Ausdrucksstärke von \mathcal{ALC} beweisen.

Das ist schwierig mit syntaktischen Argumenten,
aber Bisimulation erlaubt **semantisches Argument!**

Theorem 3.4

In \mathcal{ALC} sind **nicht ausdrückbar:**

- das \mathcal{ALCI} -Konzept $\exists r^-.T$
- die \mathcal{ALCQ} -Konzepte
 - $\leq n r.T$, für alle $n > 0$
 - $\geq n r.T$, für alle $n > 1$

T 3.3

Ausdrucksstärke und Bisimulation

Im Beweis von Theorem 3.4 gesehen:

Die Argumentation zum Beweis der Nicht-Ausdrückbarkeit läuft immer auf dasselbe hinaus:

Theorem 3.5

Sei E eine Eigenschaft. Wenn es Interpretationen $\mathcal{I}_1, \mathcal{I}_2$ und Elemente $d_1 \in \Delta^{\mathcal{I}_1}$ und $d_2 \in \Delta^{\mathcal{I}_2}$ gibt, so dass

- $(\mathcal{I}_1, d_1) \in E$ und $(\mathcal{I}_2, d_2) \notin E$ sowie
- $(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$,

dann ist E nicht in \mathcal{ALC} ausdrückbar.

Baummodelleigenschaft

Nachweis der Baummodelleigenschaft

- Interpretation \mathcal{I} ist **Baum**, wenn $(\Delta^{\mathcal{I}}, \bigcup_r r^{\mathcal{I}})$ Baum ist.
(endlich oder unendlich)
- Konzept C und TBox \mathcal{T} haben **gemeinsames Baummodell** \mathcal{I} ,
wenn \mathcal{I} Modell von \mathcal{T} und Baum mit Wurzel d sowie $d \in C^{\mathcal{I}}$.

\mathcal{ALC} hat die Baummodelleigenschaft:

Theorem 3.6

Wenn ein \mathcal{ALC} -Konzept C bzgl. einer \mathcal{ALC} -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames Baummodell \mathcal{I} .

T 3.4

Ist Grundlage für Algorithmen, erlaubt Aussagen über Ausdrucksstärke.

Beweis mittels fundamentaler Modellkonstruktion: **Unravelling**.

Unravelling

Sei \mathcal{I} eine Interpretation und $d \in \Delta^{\mathcal{I}}$.

d -Pfad in \mathcal{I} : Sequenz $d_1 d_2 \cdots d_n$, $n \geq 1$, mit $d_1 = d$ und
für alle $i < n$: es gibt Rollenname r mit $(d_i, d_{i+1}) \in r^{\mathcal{I}}$ **T 3.5**

$\text{end}(d_1 d_2 \cdots d_n) = d_n$

Definition 3.7 (Unravelling)

Unravelling von \mathcal{I} an Stelle d ist folgende Interpretation \mathcal{J} :

$\Delta^{\mathcal{J}} =$ Menge aller d -Pfade in \mathcal{I}

$A^{\mathcal{J}} = \{p \in \Delta^{\mathcal{J}} \mid \text{end}(p) \in A^{\mathcal{I}}\}$

$r^{\mathcal{J}} = \{(p, p') \in \Delta^{\mathcal{J}} \times \Delta^{\mathcal{J}} \mid$
es gibt $e : p' = p \cdot e$ und $(\text{end}(p), e) \in r^{\mathcal{I}}\}$

für alle Konzeptnamen A und Rollenamen r .

T 3.5 Forts.

Baummodelleigenschaft und Bisimulation

Sei \mathcal{J} Unravelling von \mathcal{I} an Stelle d .

Lemma 3.8

Für alle \mathcal{ALC} -Konzepte C und alle $p \in \Delta^{\mathcal{J}}$ gilt:

$$\text{end}(p) \in C^{\mathcal{I}} \quad \text{gdw.} \quad p \in C^{\mathcal{J}}$$

T 3.6

Es folgt Theorem 3.6:

Theorem 3.6 (wiederholt)

Wenn ein \mathcal{ALC} -Konzept C bzgl. einer \mathcal{ALC} -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames Baummodell \mathcal{I} .

T 3.7

Baummodelleigenschaft gilt **nicht** in Prädikatenlogik 1. Stufe (FO).

Bisimulation versus Ausdrucksstärke

Entsprechen Bisimulationen *genau* der Ausdrucksstärke von \mathcal{ALC} ?

Nein!

Gegenrichtung von Thm. 3.2 (also auch von Thm. 3.5) gilt nicht:

Es gibt Interpretationen \mathcal{I} und \mathcal{J} und $d \in \mathcal{I}$, $e \in \mathcal{J}$, so dass

- $d \in C^{\mathcal{I}}$ gdw. $e \in C^{\mathcal{J}}$ für alle \mathcal{ALC} -Konzepte C ,
- aber $(\mathcal{I}, d) \not\sim (\mathcal{J}, e)$.

T 3.8

Sie gilt allerdings für bestimmte Klassen von Interpretationen, z. B.:

- die Klasse aller endlichen Interpretationen
- die Klasse aller Interpretationen mit endlicher Verzweigungszahl
-

Solche Klassen nennt man **Hennessy-Milner-Klassen**.

Bisimulation für Erweiterungen von \mathcal{ALC}

Für \mathcal{ALCI} , \mathcal{ALCQ} , \mathcal{ALCQI} gibt es auch Bisimulationsbegriffe.

Zusätzliche Bedingungen für \mathcal{ALCI} :

- 4 Wenn $d_1 \rho d_2$ und $(d'_1, d_1) \in r^{\mathcal{I}_1}$ für beliebigen Rollennamen r , dann gibt es ein $d'_2 \in \Delta^{\mathcal{I}_2}$ mit $d'_1 \rho d'_2$ und $(d'_2, d_2) \in r^{\mathcal{I}_2}$.
- 5 Wenn $d_1 \rho d_2$ und $(d'_2, d_2) \in r^{\mathcal{I}_2}$ für beliebigen Rollennamen r , dann gibt es ein $d'_1 \in \Delta^{\mathcal{I}_1}$ mit $d'_1 \rho d'_2$ und $(d'_1, d_1) \in r^{\mathcal{I}_1}$.

Dann gilt Theorem 3.2 und man kann z. B.

eine Variante der Baummodelleigenschaft für \mathcal{ALCI} beweisen.

(Kanten im Baum können auch zur Wurzel gerichtet sein)

Auch \mathcal{ALCQ} und \mathcal{ALCQI} haben die Baummodelleigenschaft.

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

1 Bisimulation

2 Endliche Modelleigenschaft und Filtration

Überblick

Endliche Modelleigenschaft: wenn C erfüllbar bzgl. \mathcal{T} ,
dann haben C und \mathcal{T} auch *endliches* gemeinsames Modell.

Filtration wandelt Modell von Konzept und TBox in endliches Modell.

Zentrale Idee: Identifikation **ununterscheidbarer** Elemente
(im Sinne von: erfüllen dieselben Konzepte)

Größe von Konzepten und TBoxen

Definition 3.9

Die **Größe** $|C|$ eines \mathcal{ALC} -Konzeptes C ist induktiv definiert:

$$|A| = 1$$

$$|\neg C| = |C| + 1$$

$$|C \sqcap D| = |C \sqcup D| = |C| + |D| + 1$$

$$|\exists r.C| = |\forall r.C| = |C| + 3$$

Die **Größe** $|\mathcal{T}|$ einer TBox \mathcal{T} ist:

$$\sum_{C \sqsubseteq D \in \mathcal{T}} |C| + |D| + 1$$

Intuitiv: Anzahl Symbole in C bzw. \mathcal{T}

Endliche/beschränkte Modelleigenschaft

\mathcal{ALC} hat **endliche Modelleigenschaft**:

Theorem 3.10

Wenn ein \mathcal{ALC} -Konzept C bzgl. einer \mathcal{ALC} -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames **endliches** Modell.

\mathcal{ALC} hat sogar **beschränkte Modelleigenschaft**:

Theorem 3.11

Wenn ein \mathcal{ALC} -Konzept C bzgl. einer \mathcal{ALC} -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames Modell **der Kardinalität**
 $\leq 2^{|\mathcal{C}|+|\mathcal{T}|}$.

Typen: Intuition

Im Folgenden sei C \mathcal{ALC} -Konzept und \mathcal{T} TBox, so dass C erfüllbar bzgl. \mathcal{T} .

Wir benutzen Typen. Ein Typ ...

- ist Menge von Konzepten
- beschreibt ein Element $d \in \Delta^{\mathcal{I}}$ in einer Interpretation \mathcal{I}
- ist eingeschränkt auf **Teilkonzepte** von C und \mathcal{T} (um Endlichkeit zu erreichen)

Typen sind zentral für viele Techniken in Beschreibungslogik.

Teilkonzepte

Definition 3.12 (Teilkonzepte)

- $\text{sub}(C)$ ist Menge der Teilkonzepte von C , einschließlich C .
- $\text{sub}(\mathcal{T}) := \bigcup_{C \sqsubseteq D \in \mathcal{T}} \text{sub}(C) \cup \text{sub}(D)$
- $\text{sub}(C, \mathcal{T}) := \text{sub}(C) \cup \text{sub}(\mathcal{T})$

T 3.9

Lemma 3.13

$$|\text{sub}(C, \mathcal{T})| \leq |C| + |\mathcal{T}|$$

Beweis: per Induktion über C (siehe Übungen)

Typen: Definition

Definition 3.14 (Typ von d)

Sei \mathcal{I} Interpretation, $d \in \Delta^{\mathcal{I}}$. Der **Typ** $t_{\mathcal{I}}(d)$ von d in \mathcal{I} ist

$$t_{\mathcal{I}}(d) = \{D \in \text{sub}(C, \mathcal{T}) \mid d \in D^{\mathcal{I}}\}.$$

T 3.10

Lemma 3.15

Für jede Interpretation \mathcal{I} gilt: $\#\{t_{\mathcal{I}}(d) \mid d \in \Delta^{\mathcal{I}}\} \leq 2^{|C|+|\mathcal{T}|}$

(direkte Konsequenz aus Lemma 3.13)

Filtration: Idee

- Gegeben Interpretation \mathcal{I} ,
identifiziere alle Elemente gleichen Typs.
- Danach kommt also jeder Typ nur einmal vor.
- Nach Lemma 3.15 gibt es nur $\leq 2^{|\mathcal{C}|+|\mathcal{T}|}$ viele Typen.
- Wenn \mathcal{I} Modell von \mathcal{C} und \mathcal{T} , so auch das Resultat.

Filtration

Definition 3.16 (Filtration)

Sei \mathcal{I} Interpretation. Definiere Äquivalenzrelation \sim auf $\Delta^{\mathcal{I}}$:

$$d \sim e \quad \text{gdw.} \quad t_{\mathcal{I}}(d) = t_{\mathcal{I}}(e)$$

Wir bezeichnen die Äquivalenzklasse von $d \in \Delta^{\mathcal{I}}$ bzgl. \sim mit $[d]$.

Die **Filtration von \mathcal{I} bzgl. C und \mathcal{T}** ist folgende Interpretation \mathcal{J} :

$$\Delta^{\mathcal{J}} = \{[d] \mid d \in \Delta^{\mathcal{I}}\}$$

$$A^{\mathcal{J}} = \{[d] \mid d \in A^{\mathcal{I}}\} \quad \text{für alle } A \in \text{sub}(C, \mathcal{T})$$

$$r^{\mathcal{J}} = \{([d], [e]) \mid \exists d' \in [d], e' \in [e] : (d', e') \in r^{\mathcal{I}}\}$$

für alle Rollennamen r

($A^{\mathcal{J}}$ ist wohldefiniert: Repräsentantenunabhängigkeit)

T 3.10 Forts.

Theorem 3.17

Wenn \mathcal{I} Modell von C und \mathcal{T} , so auch \mathcal{J} .

T 3.11

Endliche/beschränkte Modelleigenschaft

Theorem 3.11 folgt nun unmittelbar aus Thm. 3.17 und Lem. 3.15.

Theorem 3.11 (wiederholt)

Wenn ein ALC -Konzept C bzgl. einer ALC -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames Modell der Kardinalität $\leq 2^{|\mathcal{C}|+|\mathcal{T}|}$.

Ähnliche Resultate lassen sich für $ALCI$ und $ALCQ$ beweisen
(mit derselben Schranke).

Endliche/beschränkte Modelleigenschaft

Im Kontrast zu ALC :

Theorem 3.18

$ALCQI$ hat **nicht** die endliche Modelleigenschaft.

Beweis:

A hat bzgl. folgender TBox nur unendliche Modelle:

$$T \sqsubseteq \exists r. \neg A$$

$$T \sqsubseteq \leq 1 r^{-}.T$$

T 3.12

Filtration ist also nicht für $ALCQI$ anwendbar.

Entscheidbarkeit

Theorem 3.11 (wiederholt)

Wenn ein \mathcal{ALC} -Konzept C bzgl. einer \mathcal{ALC} -TBox \mathcal{T} erfüllbar ist, dann haben C und \mathcal{T} ein gemeinsames Modell der Kardinalität $\leq 2^{|C|+|\mathcal{T}|}$.

Daraus folgt, dass Erfüllbarkeit entscheidbar ist:

Algorithmus für Erfüllbarkeit:

Gegeben C und \mathcal{T} , so dass $|C| + |\mathcal{T}| = n$,

- Erzeuge alle Interpretation \mathcal{I} mit $|\Delta|^{\mathcal{I}} \leq 2^n$
(es gibt „nur“ höchstens $2^{2^{5n}}$ viele davon).
- Überprüfe, ob \mathcal{I} Modell von C und \mathcal{T} ist
(in Zeit polynomiell in \mathcal{I} , C und \mathcal{T}).

T 3.13

Entscheidbarkeit

Lemma 3.19

Gegeben sei \mathcal{ALC} -Konzept C und endl. Interpretation \mathcal{I} .

Man kann in **polynomieller Zeit**

— genauer: in Zeit $\mathcal{O}(|C| \cdot |\Delta^{\mathcal{I}}|^2)$ —

die Extension $C^{\mathcal{I}}$ berechnen.

Korollar 3.20

Gegeben seien C, \mathcal{T} in \mathcal{ALC} und endl. Interpretation \mathcal{I} .

Man kann in **polynomieller Zeit**

— genauer: in Zeit $\mathcal{O}((|\mathcal{T}| + |C|) \cdot |\Delta^{\mathcal{I}}|^2)$ —

entscheiden, ob \mathcal{I} ein Modell von C **und** \mathcal{T} ist.

(d. h.: *Model checking* ist in Polynomialzeit)

Entscheidbarkeit

Beweis von Lemma 3.19.

Folgender rekursiver Algorithmus berechnet Extension von C in \mathcal{I} :

```

procedure ext( $C, \mathcal{I}$ )
case  $C = A$ :      return  $A^{\mathcal{I}}$ 
   $C = \neg D$ :     return  $\Delta^{\mathcal{I}} \setminus \text{ext}(D, \mathcal{I})$ 
   $C = D \sqcap E$ :  return  $\text{ext}(D, \mathcal{I}) \cap \text{ext}(E, \mathcal{I})$ 
   $C = D \sqcup E$ :  return  $\text{ext}(D, \mathcal{I}) \cup \text{ext}(E, \mathcal{I})$ 
   $C = \exists r.D$ :   return  $\{d \in \Delta^{\mathcal{I}} \mid \exists e \in \text{ext}(D, \mathcal{I}) : (d, e) \in r^{\mathcal{I}}\}$ 
   $C = \forall r.D$ :   return  $\{d \in \Delta^{\mathcal{I}} \mid \forall e \text{ mit } (d, e) \in r^{\mathcal{I}} : e \in \text{ext}(D, \mathcal{I})\}$ 
endcase
  
```

Zeitaufwand dieses Algorithmus ist $\mathcal{O}(|C| \cdot |\Delta^{\mathcal{I}}|^2)$:

- Anzahl der (rekursiven) Aufrufe = $|\text{sub}(C)| \leq |C|$
- pro Aufruf Zeitaufwand $\mathcal{O}(|\Delta^{\mathcal{I}}|^2)$:
simple Operationen auf Teilmengen von $\Delta^{\mathcal{I}}$

Entscheidbarkeit

Theorem 3.21

In \mathcal{ALC} ist Erfüllbarkeit bzgl. TBoxen entscheidbar.

Komplexität: 2ExpTime (doppelt exponentielle Zeit)

- 2-exponentiell viele Interpretationen müssen geprüft werden.
- Jede Prüfung braucht polynomielle Zeit.

Dieser Ansatz ist **kaum tauglich für die Praxis:**

- Die Komplexität ist höher als nötig.
- Das Aufzählen aller Modelle ist nicht praktikabel.

Literatur für dieses Kapitel



F. Baader, I. Horrocks, C. Lutz, U. Sattler.

An Introduction to Description Logic.

Cambridge University Press, 2017.

Kapitel 3: A Little Bit of Model Theory

In SUUB verfügbar: <https://tinyurl.com/suub-intro-dl-ebook>

<https://tinyurl.com/suub-intro-dl>