Kapitel 3

Ausdrucksstärke
und Modellkonstruktionen

Ziel des Kapitels

Die wichtigsten Eigenschaften einer BL sind

- Ausdrucksstärke und
- Komplexität

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

Wir

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

Ausdrucksstärke
und Modellkonstruktionen

Bisimulation

Überblick

- Bisimulation ist graphentheoretischer Begriff
  beschreibt die "Ähnlichkeit" zwischen Graphen

- Hängt sehr eng zusammen mit der Ausdrucksstärke von $\mathcal{ALC}$

- Wir formalisieren den Zusammenhang, betrachten drei Anwendungen:

  - Beweis der Nicht-Ausdrückbarkeit von Eigenschaften

  - Baummodelleigenschaft

  - Disjunkte Vereinigungen

Bisimulation

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

- $d_1 \; \rho \; d_2$ impliziert dass $d_1 \in A^{\mathcal{I}_1}$ gdw. $d_2 \in A^{\mathcal{I}_2}$, für alle $A \in \mathsf{N_C}$
- $d_1 \; \rho \; d_2$ und $(d_1, d_1') \in r^{\mathcal{I}_1}$ impliziert die Existenz eines $d_2' \in \Delta^{\mathcal{I}_2}$ mit $d_1' \; \rho \; d_2'$ und $(d_2, d_2') \in r^{\mathcal{I}_1}$, für alle $r \in \mathsf{N_R}$
- $d_1 \; \rho \; d_2$ und $(d_2, d_2') \in r^{\mathcal{I}_2}$ impliziert die Existenz eines $d_1' \in \Delta^{\mathcal{I}_1}$ mit $d_1' \; \rho \; d_2'$ und $(d_1, d_1') \in r^{\mathcal{I}_2}$, für alle $r \in \mathsf{N_R}$

T3.1

Bisimulation

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

$(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$: 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:
$d_1 \in C^{\mathcal{I}_1}$ gdw. $d_2 \in C^{\mathcal{I}_2}$.

T3.2

Intuitiv: $\mathcal{ALC}$ kann nicht zwischen d_1 und d_2 "unterscheiden".

Ausdrucksstärke

Wir interessieren uns für Eigenschaften von Element d in Interpretation $\mathcal{I}$:

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

Frage: ist Eigenschaft *ausdrückbar* in $\mathcal{ALC}$ (durch Konzept beschreibbar)?

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

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

Ausdrückbarkeit: gibt es äquivalentes $\mathcal{ALC}$-Konzept?

Anwendungen von Bisimulation I

Beschränkungen der Ausdrucksstärke von $\mathcal{ALC}$ beweisen.

Das ist schwierig mit syntaktischen Argumenten,

Bisimulation erlaubt semantisches Argument!

Theorem 3.3.
In $\mathcal{ALC}$ sind nicht ausdrückbar:

- das $\mathcal{ALCI}$-Konzept $\exists r^-.\top$;
- die $\mathcal{ALCQ}$-Konzepte
  - $(\leqslant n\ r\ \top)$ für alle $n > 0$ und
  - $(\geqslant n\ r\ \top)$, für alle $n > 1$.

T3.3

Anwendungen von Bisimulation II

$\mathcal{ALC}$ hat die *Baummodelleigenschaft*:

Theorem 3.4.
Wenn ein $\mathcal{ALC}$-Konzept C bzgl. einer $\mathcal{ALC}$-TBox $\mathcal{T}$ erfüllbar ist, dann haben C und $\mathcal{T}$ ein gemeinsames Baummodell $\mathcal{I}$.

$(\Delta^{\mathcal{I}}, \bigcup_{r \in \mathsf{N_R}} r^{\mathcal{I}})$ Baum (endl. oder unendl.), Wurzel in $C^{\mathcal{I}}$.

Grundlage für Algorithmen, erlaubt Aussagen über Ausdrucksstärke

Nicht ausdrückbar als Konzepte z.B.:

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

Beweis mittels fundamentaler Modellkonstruktion: Unravelling.

Anwendungen von Bisimulation II

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

d-Pfad in $\mathcal{I}$: Sequenz $d_0 d_1 \cdots d_{n-1}$, $n \geq 1$, mit

- $d_0 = d$
- für alle $i < n$: es gibt $r \in \mathsf{N_R}$ mit $(d_i, d_{i+1}) \in r^{\mathcal{I}}$

Wir setzen $\mathsf{end}(d_0 \cdots d_{n-1}) = d_{n-1}$

T3.4

Definition 3.5. (Unravelling)
Unravelling von $\mathcal{I}$ an Stelle d ist folgende Interpretation $\mathcal{J}$:

$$
\begin{aligned}
\Delta^{\mathcal{J}} &= \text{Menge aller } d\text{-Pfade in } \mathcal{I} \\
A^{\mathcal{J}} &= \{p \in \Delta^{\mathcal{I}} \mid \mathsf{end}(p) \in A^{\mathcal{I}}\} \\
r^{\mathcal{J}} &= \{(p, p') \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid \exists e : p' = p \cdot e \text{ und } (\mathsf{end}(p), e) \in r^{\mathcal{I}}\}
\end{aligned}
$$

für alle $A \in \mathsf{N_C}$ und $r \in \mathsf{N_R}$.

T3.4 cont

Anwendungen von Bisimulation II

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

Lemma 3.6.
Für alle $\mathcal{ALC}$-Konzepte C und alle $p \in \Delta^{\mathcal{J}}$ gilt:
end$(p) \in C^{\mathcal{I}}$ gdw $p \in C^{\mathcal{J}}$.

T3.5

Es folgt Theorem 3.4.

T3.6

Baummodelleigenschaft gilt nicht in FO.

Anwendungen von Bisimulation III

Definition 3.7. (Disjunkte Vereinigung)
Seien $\mathcal{I}$ und $\mathcal{J}$ Interpretationen mit $\Delta^{\mathcal{I}} \cap \Delta^{\mathcal{J}} = \emptyset$.
Die disjunkte Vereinigung $\mathcal{I} \uplus \mathcal{J}$ von $\mathcal{I}$ und $\mathcal{J}$ ist definiert wie folgt:

$$\begin{aligned}
\Delta^{\mathcal{I} \uplus \mathcal{J}} &= \Delta^{\mathcal{I}} \cup \Delta^{\mathcal{J}} \\
A^{\mathcal{I} \uplus \mathcal{J}} &= A^{\mathcal{I}} \cup A^{\mathcal{J}} \text{ für alle } A \in \mathsf{N_C} \\
r^{\mathcal{I} \uplus \mathcal{J}} &= r^{\mathcal{I}} \cup r^{\mathcal{J}} \text{ für alle } r \in \mathsf{N_R}
\end{aligned}$$

Lemma 3.8.
Für alle $\mathcal{ALC}$-Konzepte C und alle $d \in \Delta^{\mathcal{I}}$ gilt: $d \in C^{\mathcal{I}}$ gdw. $d \in C^{\mathcal{I} \uplus \mathcal{J}}$.

T3.7

Korollar 3.8b.
Wenn $\mathcal{I}$ Modell von Konzept C (bzw. TBox $\mathcal{T}$), dann auch
$\mathcal{I} \uplus \mathcal{J}$ Modell von C (bzw. $\mathcal{T}$).

Man kann z.B. zeigen, dass $\exists y.x \neq y$ nicht ausdrückbar.

Bisimulation

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

Nein! Gegenrichtung von Theorem 3.2 gilt nicht:

Es gibt Interpretationen $\mathcal{I}$ und $\mathcal{J}$ und $d \in \mathcal{I}$, $x \in \mathcal{J}$ so dass
$d \in C^{\mathcal{I}}$ gdw. $x \in C^{\mathcal{J}}$ für all $\mathcal{ALC}$-Konzepte C, aber $(\mathcal{I}, d) \not\sim (\mathcal{J}, x)$.

T3.8

Sie gilt allerdings für verschiedene Modellklassen, z.B.:

- die Klasse aller endlichen Modelle;
- die Klasse aller Modelle mit endlicher Verzweigungszahl;
- die Klasse aller modal saturierten Modelle.

Derartige Klassen nennt man Henessy-Milner Klassen.

Erweiterungen von ALC

Für $\mathcal{ALCI}$ und $\mathcal{ALCQI}$ gibt es ebenfalls Bisimulationsbegriffe.

T3.9

Zusätzliche Bedingungen für $\mathcal{ALCI}$:

- $d_1 \; \rho \; d_2$ und $(d_1', d_1) \in r^{\mathcal{I}_1}$ impliziert die Existenz eines $d_2' \in \Delta^{\mathcal{I}_2}$ mit $d_1' \; \rho \; d_2'$ und $(d_2', d_2) \in r^{\mathcal{I}_1}$, für alle $r \in \mathsf{N_R}$

- $d_1 \; \rho \; d_2$ und $(d_2', d_2) \in r^{\mathcal{I}_2}$ impliziert die Existenz eines $d_1' \in \Delta^{\mathcal{I}_1}$ mit $d_1' \; \rho \; d_2'$ und $(d_1', d_1) \in r^{\mathcal{I}_2}$, für alle $r \in \mathsf{N_R}$

Dann gilt Theorem 3.2 und man kann z.B. Baummodell-Eigenschaft für $\mathcal{ALCI}$ beweisen (Kanten im Baum können auch zur Wurzel gerichtet sein).

Auch Lemma 3.8 gilt für $\mathcal{ALCI}$.

Erweiterungen von ALC

Auch für $\mathcal{ALCQ}$ kann man Bisimulationsbegriff finden.

Für $\mathcal{ALCQ}$ gelten sowohl Baummodelleigenschaft als auch Lemma 3.8

Ebenso für $\mathcal{ALCQI}$

Ausdrucksstärke und Modellkonstruktionen

Endliche Modelleigenschaft und Filtration

Überblick

- Filtration ist wichtige Methode aus der Modallogik
- Wandelt Modell eines Konzeptes und einer TBox in Modell kleinerer Größe;
- Zentrale Idee: Identifikation ununterscheidbarer Elemente (Bisimulation!)
- Erlaubt es, die Existenz endlicher/kleiner Modelle zu beweisen

Größe von Konzepten und TBoxen

Definition 3.9.
Größe $|C|$ eines $\mathcal{ALC}$-Konzeptes C ist induktiv definiert:

$$
\begin{aligned}
|A| &= 1 \\
|\neg C| &= |C| + 1 \\
|C \sqcap D| &= |C \sqcup D| = |C| + |D| + 1 \\
|\exists r.C| &= |\forall r.C| = |C| + 3
\end{aligned}
$$

Größe $|\mathcal{T}|$ einer (generellen oder azyklischen) TBox $\mathcal{T}$ ist

$$
\sum_{C \sqsubseteq D \in \mathcal{T}} |C| + |D| + 1 + \sum_{A \equiv C \in \mathcal{T}} |C| + 2
$$

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.

Dies gilt nicht in FO.

$\mathcal{ALC}$ 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 $\mathbf{2}^{|C|+|\mathcal{T}|}$.

Typ

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

Im folgenden definieren wir den Begriffs eines "Typs":

- ist Menge von Konzepten;
- beschreibt einen Punkt $d \in \Delta^{\mathcal{I}}$ in einer Interpretation $\mathcal{I}$;
- um Endlichkeit zu erreichen: Einschränkung auf endliche Konzeptmenge:

  Konzepte, die für das "Modell-sein" von C und $\mathcal{T}$ relevant sind.

Zentral für viele Techniken im Bereich Beschreibungs-/Modallogik

Typ

Definition 3.12. (Teilkonzepte)

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

T3.10

Lemma 3.13.
$|\mathsf{sub}(C, \mathcal{T})| \leq |C| + |\mathcal{T}|$.

Für später: da jede azyklische TBox als generelle gesehen werden kann, definiert Definition 3.12 auch $\mathsf{sub}(C, \mathcal{T})$ für azyklische TBoxen.

Typ

Definition 3.14. (Typ von $\boldsymbol{d}$)
Sei $\mathcal{I}$ Interpretation, $\boldsymbol{d} \in \boldsymbol{\Delta}^{\mathcal{I}}$. *Typ* $\boldsymbol{t_{\mathcal{I}}(d)}$ von $\boldsymbol{d}$ in $\mathcal{I}$ bzgl. ist

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

T3.11

Lemma 3.15.
Für jede Interpretation $\mathcal{I}$ gilt $|\{\boldsymbol{t_{\mathcal{I}}(d)} \mid \boldsymbol{d} \in \boldsymbol{\Delta}^{\mathcal{I}}\}| \leq 2^{|C|+|\mathcal{T}|}$.

Filtration: Idee

- Gegeben Interpretation $\mathcal{I}$, identifiziere alle Elemente gleichen Typs
- Identifizieren von $\boldsymbol{d}$ und $\boldsymbol{e}$: biege alle ein+ausgehenden Kanten von $\boldsymbol{e}$ zu $\boldsymbol{d}$ um, dann lösche $\boldsymbol{e}$
- Danach kommt also jeder Typ nur einmal vor
- Nach Lemma 3.15 gibt es nur $\mathbf{2}^{|C|+|\mathcal{T}|}$ viele Typen
- Wenn $\mathcal{I}$ Modell von $\boldsymbol{C}$ und $\mathcal{T}$, so auch das Resultat

Filtration

Definition 3.16. (Filtration)
Sei $\mathcal{I}$ Modell von C und $\mathcal{T}$. Definiere Äquivalenzrelation $\sim$ auf $\Delta^{\mathcal{I}}$:

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

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 $\mathcal{J}$:

$$\begin{aligned}
\Delta^{\mathcal{J}} &= \{[d] \mid d \in \Delta^{\mathcal{I}}\} \\
A^{\mathcal{J}} &= \{[d] \mid d \in A^{\mathcal{I}}\} \text{ für alle } A \in \mathsf{sub}(C, \mathcal{T}) \\
A^{\mathcal{J}} &= \emptyset \text{ für alle } A \in \mathsf{sub}(C, \mathcal{T}) \text{ (irrelevant)} \\
r^{\mathcal{J}} &= \{([d], [e] \mid \exists d' \in [d], e' \in [e] : (d', e') \in r^{\mathcal{I}}\} \text{ für alle } r \in \mathsf{N_R}
\end{aligned}$$

Beachte: $A^{\mathcal{J}}$ is wohldefiniert (Repräsentantenunabhängigkeit).

Theorem 3.17.
Wenn $\mathcal{I}$ Modell von C und $\mathcal{T}$, so auch $\mathcal{J}$. T3.12

Endliche/Beschränkte Modelleigenschaft.

Theorem 3.11 folgt nun unmittelbar aus Theorem 3.17 und Lemma 3.15.

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 $\mathbf{2}^{|C|+|\mathcal{T}|}$.

Ähnliche Resultate lassen sich für $\mathcal{ALCI}$ und $\mathcal{ALCQ}$ beweisen.

Endliche/Beschränkte Modelleigenschaft.

Theorem 3.18.
$\mathcal{ALCQI}$ hat nicht die endliche Modelleigenschaft.

Beweis:

A hat nur unendliche Modelle bzgl. folgender TBox:

$$\top \sqsubseteq \exists r.\neg A$$
$$\top \sqsubseteq (\leqslant 1\; r^{-}\; \top)$$

T3.13

Filtration für $\mathcal{ALCQI}$ also nicht anwendbar.