Beschreibungslogiken

Die Beschreibungslogik $\mathcal{ALC}$

Die elementaren Grundbegriffe jeder Beschreibungslogik sind

Konzept und Rolle

Eine Instanz der Beschreibungslogik $\mathcal{ALC}$ wird bestimmt durch die Festlegung

1. eines endlichen Vorrats $\mathbf{C}$ von Symbolen für Konzepte
2. und eines endlichen Vorrats $\mathbf{R}$ von Symbolen für Rollen.

Die Vereinigung $V = C \cup R$ nennen wir das Vokabular.

$\mathcal{ALC}$-Ausdrücke

Definition

$\mathcal{ALC}$-Ausdrücke über einem Vokabular $\mathbf{V} = \mathbf{C} \cup \mathbf{R}$:

1. jedes Konzeptsymbol C aus $\mathbf{C}$ ist ein $\mathcal{ALC}$-Ausdruck,
2. sind $C_1, \ldots, C_k$ $\mathcal{ALC}$-Ausdrücke, dann sind auch

   $$C_1 \sqcap \ldots \sqcap C_k, \qquad C_1 \sqcup \ldots \sqcup C_k, \qquad \neg C_1$$

   $\mathcal{ALC}$-Ausdrücke,
3. ist C ein $\mathcal{ALC}$-Ausdruck, R ein Rollensymbol aus $\mathbf{R}$, dann sind auch

   $$\exists R.C \text{ und } \forall R.C$$

   $\mathcal{ALC}$-Ausdrücke.

Interpretationen für ALC

Definition

Eine Interpretation für $\mathcal{ALC}$ über einem Vokabular $\mathbf{V}$ besteht aus einem Grundbereich von Objekten, der üblicherweise mit Δ bezeichnet wird und einer Funktion I, für die gilt:

1. für jedes Konzeptsymbol $C \in \mathbf{C}$ ist C^I eine Teilmenge Δ
2. für jedes Rollensymbol $R \in \mathbf{R}$ ist R^I eine binäre Relation aus Δ, i.e. $R^I \subseteq \Delta \times \Delta$.

Semantik für ALC

Sei I eine Interpretation.

1. $(C_1 \sqcap \ldots \sqcap C_k)^I = C_1^I \cap \ldots \cap C_k^I$
2. $(C_1 \sqcup \ldots \sqcup C_k)^I = C_1^I \cup \ldots \cup C_k^I$
3. $(\neg C)^I = \Delta \setminus C^I$
4. $(\exists R.C)^I = \{d \in \Delta \mid$ es gibt $(d, e) \in R^I$ mit $e \in C^I\}$,
5. $(\forall R.C)^I = \{d \in \Delta \mid$ für alle $(d, e) \in R^I$ gilt $e \in C^I\}$

Beispiele

Sind M, F, P Namen für Konzepte, die wir intuitiv als die Klasse aller Männer, aller Frauen bzw. aller Personen deuten und ist $R = ist\_direkter\_Vorfahr\_von$ eine Rolle, dann sind

$$Va = M \sqcap \exists R.P$$

$$Mu = F \sqcap \exists R.P$$

$$E = Va \sqcup Mu$$

korrekt geformte Formeln in $\mathcal{ALC}$, welche die Konzepte Vater, Mutter und Eltern einführen.

Übersetzt in prädikatenlogische Notation:

$$Va(x) = M(x) \land \exists y(R(x,y) \land P(y))$$

$$Mu(x) = F(x) \land \exists y(R(x,y) \land P(y))$$

$$E(x) = Va(x) \lor Mu(x)$$

Aufgabestellungen

C_1, C_2 seien Formeln einer Beschreibungslogik, so ist man daran interessiert festzustellen, ob für alle Interpretationen I

C_1^I dasselbe Konzept wie C_2^I ist	$C_1 = C_2$
C_1^I ein Teilmenge von C_2^I ist	$C_1 \sqsubseteq C_2$
C_1^I leer ist	$C_1 = \emptyset$

Reduzierte Aufgabestellungen

Wegen

$$C_1 = C_2 \text{ gdw } C_1 \sqsubseteq C_2 \wedge C_2 \sqsubseteq C_1$$

und

$$C_1 \sqsubseteq C_2 \text{ gdw } C_1 \sqcap \neg C_2 = \emptyset$$

genügt es sich auf die letzte Fragestellung $C = \emptyset?$ zu beschränken.
Die Beantwortung dieser Frage wird üblicherweise

Subsumtionsproblem

genannt.

$\mathcal{ALC}$ als multi-modale Logik

Wir übersetzen jede Formel F von $\mathcal{ALC}$ in eine Formel F^* einer multi-modalen Logik, die für jede Rolle R

modale Operatoren $\Box_R$ und $\Diamond_R$ enthält.

Bei der Übersetzung wird jedes Konzeptsymbol C durch eine aussagenlogische Variable C ersetzt.

Übersetzung

Jeder $\mathcal{ALC}$-Formel F wird die multi-modale Formel F^* wie folgt zugeordnet:

$$
\begin{array}{lcll}
F^* & = & F & \text{falls } F \text{ Konzeptsymbol} \\
(F_1 \sqcap F_2)^* & = & (F_1^* \land F_2^*) & \\
(\neg F)^* & = & \neg F^* & \\
(\forall R.F)^* & = & \Box_R F^* & \\
(\exists R.F)^* & = & \Diamond_R F^* &
\end{array}
$$

$\mathcal{K}^I$

Einer Interpretation I für $\mathcal{ALC}$ entspricht in eindeutiger Weise eine multi-modalen Kripke Stuktur $\mathcal{K}^I$.

- Die Menge der möglichen Welten von $\mathcal{K}^I$ entspricht dem Universum Δ von I.
- Jede Rolle R gibt Anlaß zu einer Zugänglichkeitsrelation R^I auf Δ, die für die Semantik von $\Box_R$ und $\Diamond_R$ zuständig ist.
- Für $d \in \Delta$ wird schließlich die Bewertung $v(d,C)$ auf *wahr* gesetzt genau dann, wenn $d \in C^I$ zutrifft.

Lemma

1. Für jede $\mathcal{ALC}$-Formel F und jede Interpretation I gilt

$$F^I = \{d \in \Delta \mid (\mathcal{K}^I, d) \models F^*\}$$

2. Es gibt eine Interpretation I und ein $d \in \Delta$ mit $d \in F^I$ genau dann, wenn F^* erfüllbar ist.
3. Es gilt $F_2 \sqsubseteq F_1$ genau, dann wenn $(F_1 \sqcap \neg F_2)^*$ unerfüllbar ist.

Beweis

zu 1:

Der Beweis wird über den Aufbau der Formel F geführt.
Ist F ein Konzeptname, dann ergibt sich nach Definition von $\mathcal{K}^I$

$$\begin{aligned}\{d \in \Delta \mid (\mathcal{K}^I, d) \models F^*\} &= \{d \in \Delta \mid v(d, F) = \textit{wahr}\} \\ &= F^I\end{aligned}$$

Wir betrachten im Induktionsschritt nur den Fall $F = \forall R.F_1$

$$\begin{aligned}&\{d \in \Delta \mid (\mathcal{K}^I, d) \models (\forall R.F_1)^*\} &= \\ &\{d \in \Delta \mid (\mathcal{K}^I, d) \models \Box_R F_1^*\} &= \\ &\{d \in \Delta \mid \text{ für alle } e \text{ mit } R(d, e) \text{ gilt } (\mathcal{K}^I, e) \models F_1^*\} &= \\ &\{d \in \Delta \mid \text{ für alle } e \text{ mit } R(d, e) \text{ gilt } e \in F_1^I\} &= \\ &(\forall R.F_1)^I\end{aligned}$$

Beweis (Forts.)

zu 2: folgt unmittelbar aus Teil 1.

zu 3:
$F_1 \sqsubseteq F_2$ gilt genau dann, wenn für alle I und alle $d \in \Delta$ gilt

$$d \notin (F_1 \sqcap \neg F_2)^I .$$

Nach Teil 1 ist das gleichbedeutend mit

$$(K^I, d) \not\models F_1^* \wedge \neg F_2^*$$

für alle I und alle $d \in \Delta$.

Das besagt nichts anderes, als daß $(F_1 \sqcap \neg F_2)^*$ nicht erfüllbar ist.

Die Beschreibungslogik $\mathcal{SHIQ}$

Ein Vokabular für $\mathcal{SHIQ}$ besteht aus

1. einer Menge $\mathbf{C}$ von Konzeptsymbolen
2. in $\mathbf{C}$ ist stets das unverselle Konzept $\top$ enthalten.
3. einer Menge $\mathbf{R}^0$ von atomaren Rollensymbolen.
4. einer Teilmenge $\mathbf{R}^0_t \subseteq \mathbf{R}^0$ von transitiven Rollensymbolen,
5. Die Menge aller Rollensymbole $\mathbf{R}$ enthält alle atomaren Rollen und für jede atomare Rolle $R \in \mathbf{R}^0$ ihre inverse Rolle R^-.

Rollenhierarchie

Eine Rollenhierarchie $\mathcal{H}$ ist eine endliche Menge von Formeln der Art

$$R_1 \sqsubseteq R_2$$

für Rollensymbole $R_i \in \mathbf{R}$.

Alle nachfolgenden Definitionen setzten eine im voraus gewählte und festgehaltene Rollenhierarchie voraus.

$\mathcal{SHIQ}$-Ausdrücke

1. jedes Konzeptsymbol C aus $\mathbf{C}$ ist ein $\mathcal{SHIQ}$-Ausdruck,
2. das Symbol $\top$ ist ein $\mathcal{SHIQ}$-Ausdruck,
3. sind $C_1, \ldots, C_k$ $\mathcal{SHIQ}$-Ausdrücke, dann sind auch
   $C_1 \sqcap \ldots \sqcap C_k \qquad C_1 \sqcup \ldots \sqcup C_k \qquad \neg C_1$,
   $\mathcal{SHIQ}$-Ausdrücke,
4. ist C ein $\mathcal{SHIQ}$-Ausdruck, R ein Rollensymbol aus $\mathbf{R}$, dann sind
   $\exists R.C \qquad \forall R.C$
   $\mathcal{SHIQ}$-Ausdrücke,
5. ist R ein einfacher Rollenbezeichner aus $\mathbf{R}$, C ein $\mathcal{SHIQ}$-Ausdruck und n eine natürliche Zahl, dann sind
   $\leq nR.C \qquad \geq nR.C$.
   $\mathcal{SHIQ}$-Ausdrücke,
   Dabei heißt ein Rollenbezeichner R einfach, wenn $R \notin R_t^0$ gilt und überdies R keine transitive Unterrolle besitzt.

Beispiel

Zusätzlich zu den Konzepten und Rollen aus dem vorangegangenen Beispiel benutzen wir noch das Konzept W, das intuitiv für alle weiblichen Personen stehen soll.
Der Ausdruck

$$C = E \sqcap \ \geq 2R.W$$

bezeichnet die Menge aller Eltern, die mindestens zwei Töchter haben.

Interpretationen für $\mathcal{SHIQ}$

Eine Interpretation für das Vokabular $\mathbf{V}$ besteht aus einem mit Δ bezeichneten Grundbereich von Objekten und einer Funktion I mit folgenden Eigenschaften:

1. für jedes $C \in \mathbf{C}$ ist C^I eine Teilmenge von Δ,
2. für das spezielle Symbol $\top \in \mathbf{C}$ gilt $\top^I = \Delta$,
3. für jedes atomare Rollensymbol $R \in \mathbf{R}^0$ ist R^I eine binäre Relation auf Δ, i.e. $R^I \subseteq \Delta \times \Delta$,
4. für jedes atomare Rollensymbol $R \in \mathbf{R}_t^0$ ist R^I sogar eine transitive Relation auf Δ,
5. für jedes atomare Rollensymbol $R \in \mathbf{R}^0$ ist $(R^-)^I$ die inverse Relation zu R^I, i.e.

$$(d_1, d_2) \in (R^-)^I \Leftrightarrow (d_2, d_1) \in R^I$$

Rollenhierarchien

Sei $\mathcal{H}$ eine Rollenhierarchie.
Es wird im folgenden vorausgesetzt, daß alle auftretenden Interpretationen I die Rollenhierarchie respektieren, d.h. es gilt

$$R_1^I \subseteq R_2^I \text{ für jede Formel } R_1 \sqsubseteq R_2 \in \mathcal{H}.$$

Semantik von $\mathcal{S}HIQ$-Ausdrücken

Sei I eine Interpretation.

1. $(\leq nR.C)^I = \{d \in \Delta \mid \#\{e \mid (d,e) \in R^I \text{ und } e \in C^I\} \leq n\}$
2. $(\geq nR.C)^I = \{d \in \Delta \mid \#\{e \mid (d,e) \in R^I \text{ und } e \in C^I\} \geq n\}$

Grundbegriffe

1. Eine endliche Menge $\mathcal{T}$ von Formeln der Form $C_1 \sqsubseteq C_2$ für $\mathcal{SHIQ}$-Ausdrücke C_1, C_2 heißt eine Terminologie
2. Eine Interpretation I heißt ein Modell einer Terminologie $\mathcal{T}$, wenn $C_1^I \subseteq C_2^I$ für alle $C_1 \sqsubseteq C_2 \in \mathcal{T}$ gilt.
3. Ein $\mathcal{SHIQ}$-Ausdruck C ist erfüllbar bezüglich $\mathcal{T}$, wenn es ein Modell I von $\mathcal{T}$ gibt mit $C^I \neq \emptyset$
4. D subsumiert C bezüglich $\mathcal{T}$, wenn für jedes Modell I von $\mathcal{T}$ gilt: $C^I \subseteq D^I$.

Theorem

Das Erfüllbarkeits- und das Subsumptionsproblem für $\mathcal{SHIQ}$ ist entscheidbar.

Beweis siehe [Horrocks et al.]

Unentscheidbarkeit

Läßt man in Definition von $\mathcal{SHIQ}$-Ausdrücken die Einschränkung auf einfache Rollen fallen, so wird das Erfüllbarkeitsproblem unentscheidbar.

Das gilt sogar dann, wenn man nur Kardinalitätsformeln der Form
$\leq nR.\top$ und $\geq nR.\top$ zuläßt
(siehe ebenfalls [Horrocks et al.])

Beispiele aus TAMBIS

TAMBIS = Transparent Access to Multiple Bioinformatics Information Sources

```
class-def protein

class-def defined holoprotein
      subclass-of protein
      slot-constraint binds
          has-value p-group

class-def defined enzyme
      subclass-of protein
      slot-constraint catalyses
          has-value reaction

disjoint metal-ion small-molecule
```

```
class-def defined holoenzyme
      subclass-of enzyme
      slot-constraint binds
          has-value p-group

class-def defined cofactor
      subclass-of (metal-ion
          or small-molecule)

class-def defined ion
      subclass-of chemical
      slot-constraint has-charge
          has-value ( not 0)
```

Beispiele aus TAMBIS

```
slot-def lysis-of
    domain reaction
    range cv-bond

class-def defined lysis
    subclass-of reaction
    slot-constraint lysis-of
        has-value cv-bond
        value-type cv-bond

class-def defined lyase
    subclass-of protein
    slot-constraint catalyses
        has-value lysis
        value-type lysis

class-def defined hydrolysis
    subclass-of reaction
    slot-constraint hydrolysis-of
        has-value cv-bond
        value-type cv-bond

slot-def hydrolysis-of
    subslot-of lysis-of

class-def defined hydrolase
    subclass-of protein
    slot-constraint catalyses
        has-value hydrolysis
        value-type hydrolysis
```