Logik Teil 3: Mehr zur Prädikatenlogik erster Stufe

Vorlesung im Wintersemester 2010

Übersicht Teil 3

- Kapitel 3.1: Sequenzenkalkül
- Kapitel 3.2: Rekursive Aufzählbarkeit, Kompaktheit und Löwenheim-Skolem
- Kapitel 3.3: Ausdrucksstärke / Grundlagen von Ehrenfeucht-Fraïsse Spielen
- Kapitel 3.4: Ehrenfeucht-Fraïsse Spiele: Anwendungen

Mehr zur Prädikatenlogik

Kapitel 3.1: Sequenzenkalkül

Sequenzenkalkül

Wir betrachten ein Kalkül für Gültigkeit in der Prädikatenlogik

Motivation:

- rekursive Aufzählbarkeit nachweisen
- einfacher Beweis für das Kompaktheitstheorem in FO
- einfache Beweise weiterer wichtiger Resultate, insbesondere Satz von Löwenheim-Skolem

Im Prinzip könnten wir wieder Resolution verwenden
(Grundlage für Theorembeweiser der Logik erster Stufe)

Wir verwenden aber einen technisch einfacheren Ansatz:

Gentzens Sequenzenkalkül

Der Einfachheit halber verzichten wir auf das Gleichheitsprädikat

Sequenz

Definition Sequenz

Eine *Sequenz* ist ein Ausdruck der Form $\Gamma \Rightarrow \Delta$ wobei Γ, $\Delta \subseteq FO$ endliche Mengen von Sätzen sind. Wir nennen

- Γ das *Antezendenz* und
- Δ das *Sukzedenz*.

Die Sequenz $\Gamma \Rightarrow \Delta$ ist *gültig* wenn $\bigwedge \Gamma \models \bigvee \Delta$, in Worten:

jedes Modell von $\bigwedge \Gamma$ macht auch mindestens einen Satz aus Δ wahr

Ist eine Sequenz $\Gamma \Rightarrow \Delta$ gültig, so schreiben wir $\models \Gamma \Rightarrow \Delta$.

Beispiele für gültige Sequenzen:

- $\{\forall x.P(x), Q(c)\} \Rightarrow \{ P(c) \wedge Q(c), R(c,d)\}$
- jede Sequenz $\Gamma \Rightarrow \Delta$ mit $\Gamma \cap \Delta \neq \emptyset$

Sequenzenkalkül

Das Sequenzenkalkül erlaubt, alle gültigen Sequenzen abzuleiten

Offensichtlich:

- FO-Satz φ ist Tautologie gdw. die Sequenz $\emptyset \Rightarrow \{\varphi\}$ gültig ist
- FO-Satz φ ist unerfüllbar gdw. die Sequenz $\{\varphi\} \Rightarrow \emptyset$ gültig ist (denn $\bigvee \emptyset$ ist unerfüllbar)

Man kann das Sequenzenkalkül also auch als Kalkül zum Ableiten aller Tautologien/unerfüllbaren *Formeln* ansehen.

Sequenzenkalkül

Die zentralen Bestandteile des SK:

- Axiome

  Sequenzen, die man ohne Beweis / Herleitung als gültig voraussetzt

- Schlussregeln

  Im Gegensatz zu Resolution/Hilbert hat das SK recht viele davon:

  2 Stück pro Operator $\neg$, $\wedge$, $\vee$, $\forall$, $\exists$,

  jeweils für die linke und die rechte Seite von Sequenzen

  (positive und negative Form der Regel)

Sequenzenkalkül

Zum Hervorheben von Formeln in Sequenzen schreiben wir

$$\Gamma, \varphi \Rightarrow \Delta, \psi \ \text{ statt } \ \Gamma \cup \{\varphi\} \Rightarrow \Delta \cup \{\psi\}$$

Definition Axiome SK

Die *Axiome* des Sequenzenkalküls (SK) sind alle Sequenzen der Form $\Gamma, \varphi \Rightarrow \Delta, \varphi$.

Axiome sind offensichtlich gültige Sequenzen

Sequenzenkalkül

$$(\neg \Rightarrow)\ \frac{\Gamma \Rightarrow \Delta, \varphi}{\Gamma, \neg\varphi \Rightarrow \Delta} \qquad (\Rightarrow \neg)\ \frac{\Gamma, \varphi \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg\varphi}$$

$$(\vee \Rightarrow)\ \frac{\Gamma, \varphi \Rightarrow \Delta \qquad \Gamma, \psi \Rightarrow \Delta}{\Gamma, \varphi \vee \psi \Rightarrow \Delta} \qquad (\Rightarrow \vee)\ \frac{\Gamma \Rightarrow \Delta, \varphi, \psi}{\Gamma \Rightarrow \Delta, \varphi \vee \psi}$$

$$(\wedge \Rightarrow)\ \frac{\Gamma, \varphi, \psi \Rightarrow \Delta}{\Gamma, \varphi \wedge \psi \Rightarrow \Delta} \qquad (\Rightarrow \wedge)\ \frac{\Gamma \Rightarrow \Delta, \varphi \qquad \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \varphi \wedge \psi}$$

$$(\exists \Rightarrow)\ \frac{\Gamma, \varphi[c] \Rightarrow \Delta}{\Gamma, \exists x.\varphi(x) \Rightarrow \Delta}$$ c nicht in $\Gamma, \Delta, \varphi(x)$

$$(\Rightarrow \exists)\ \frac{\Gamma \Rightarrow \Delta, \varphi[t]}{\Gamma \Rightarrow \Delta, \exists x.\varphi(x)}$$

$$(\forall \Rightarrow)\ \frac{\Gamma, \varphi[t] \Rightarrow \Delta}{\Gamma, \forall x.\varphi(x) \Rightarrow \Delta}$$

$$(\Rightarrow \forall)\ \frac{\Gamma \Rightarrow \Delta, \varphi[c]}{\Gamma \Rightarrow \Delta, \forall x.\varphi(x)}$$ c nicht in $\Gamma, \Delta, \varphi(x)$

Sequenzenkalkül

Definition ableitbar

Die Menge der *ableitbaren* Sequenzen ist die kleinste Menge von Sequenzen, die

- alle Axiome des SK enthält und
- abgeschlossen ist unter Regelanwendung: wenn Instanzen der Sequenzen in der oberen Zeile einer Schlussregel enthalten sind, so auch die entsprechende Instanz der unteren Zeile

Ist eine Sequenz $\Gamma \Rightarrow \Delta$ ableitbar, so schreiben wir $\vdash \Gamma \Rightarrow \Delta$.

Instanz bedeutet: $\Gamma, \Delta, \varphi, \psi$ durch konkrete Formeln/Formelmengen ersetzen

Beispiel

Sequenzenkalkül

Definition SK-Beweis

Ein *SK-Beweis* ist ein Baum, dessen Knoten auf folgende Weise mit Sequenzen beschriftet sind:

- Jedes Blatt ist mit einem Axiom beschriftet
- Jeder innere Knoten ist mit einer Instanz der unteren Zeile einer Schlussregel beschriftet
- die Kinder dieses Knotens sind dann genau mit den entsprechenden Instanzen der Sequenzen in der oberen Zeile der Regel beschriftet.

Beachte:

- jeder innere Knoten hat ein oder zwei Kinder
- eine Sequenz ist ableitbar gdw. sie als Knotenbeschriftung in einem SK-Beweis auftritt.

Beispiel

Sequenzenkalkül

Zur Erinnerung:

In der Sequenz Γ, φ darf Γ auch φ enthalten, muss aber nicht

Darum darf man bei Anwendung von $(\Rightarrow \exists)$ und $(\forall \Rightarrow)$ im SK-Beweis die verwendete Teilformel “behalten”:

Beispiel $(\forall \Rightarrow)$: $$(\forall \Rightarrow)\ \frac{\Gamma, \varphi[t] \Rightarrow \Delta}{\Gamma, \forall x.\varphi(x) \Rightarrow \Delta}$$

$$\frac{P(c) \Rightarrow P(c) \wedge P(d)}{\forall x.P(x) \Rightarrow P(c) \wedge P(d)} \qquad \frac{\color{red}{\forall x.P(x),}P(c) \Rightarrow P(c) \wedge P(d)}{\forall x.P(x) \Rightarrow P(c) \wedge P(d)}$$

Das gilt im Prinzip für alle Regeln ist, aber nur bei $(\Rightarrow \exists)$ und $(\forall \Rightarrow)$ nützlich (und notwendig!)

Korrektheit

Theorem (Korrektheit SK)

Wenn $\vdash \Gamma \Rightarrow \Delta$, dann $\models \Gamma \Rightarrow \Delta$ (jede ableitbare Sequenz ist gültig).

Beweis:

Es reicht, zu zeigen:

1. alle SK-Axiome sind gültig

   offensichtlich gilt $\bigwedge \Gamma \models \bigvee \Delta$ wenn es $\varphi \in \Gamma \cap \Delta$ gibt

2. wenn eine Sequenz $\Gamma \Rightarrow \Delta$ durch das Anwenden einer Schluss-regel auf gültige Sequenzen entsteht, dann ist $\Gamma \Rightarrow \Delta$ gültig.

   Fallunterscheidung: ein Fall pro Regel.

Vollständigkeit

Theorem (Vollständigkeit SK)

Wenn $\models \Gamma \Rightarrow \Delta$, dann $\vdash \Gamma \Rightarrow \Delta$ (jede gültige Sequenz ist ableitbar).

Beweisstrategie:

Zeige: wenn Sequenz $\Gamma \Rightarrow \Delta$ nicht ableitbar,

dann gibt es Modell $\mathfrak{A}$ für $\Gamma \cup \neg\Delta$, wobei $\neg\Delta = \{\neg\varphi \mid \varphi \in \Delta\}$

Im Prinzip möchten wir $\mathfrak{A}$ einfach aus Γ ”ablesen”,

die nicht-Ableitbarkeit von $\Gamma \Rightarrow \Delta$ soll sicherstellen, dass $\mathfrak{A} \models \neg\Delta$

Um das ”Ablesen” zu ermöglichen, muß Γ erst vervollständigt werden

denn z.B. $\Gamma = \{Q_1(c) \vee Q_2(c), \exists x.P(x)\}$ entspricht keinem Modell

Vollständigkeit

Für den Rest des Beweises fixiere Signatur $\tau = \mathsf{Sig}(\Gamma \cup \Delta) \cup C$,

C abzählbar unendliche Menge neuer Konstantensymbole

Wir arbeiten o.B.d.A. mit reduzierten Formeln (nur $\neg$, $\wedge$, $\exists$):

Eine Sequenz $\Gamma \Rightarrow \Delta$ ist *reduziert* wenn alle Sätze in Γ, Δ reduziert sind.

Lemma

Wenn alle gültigen reduzierten Sequenzen ableitbar sind,
dann sind alle gültigen Sequenzen ableitbar.

Vollständigkeit

Definition (Herbrandstruktur)

Eine τ-Struktur $\mathfrak{H}$ heisst *Herbrandstruktur* wenn

- ihr Universum H die Menge aller Grundterme der Signatur τ ist
- alle Funktionssymbole
  $f \in F^n(\tau)$, $n \geq 0$, werden durch syntaktisches Anwenden interpretiert:

$$f^{\mathfrak{H}}(t_1, \ldots, t_n) = f(t_1, \ldots, t_n)$$

Beispiel

Beachte: Für alle Konstanten c gilt also $c^{\mathfrak{H}} = c$, für alle Grundterme $t^{\mathfrak{H}} = t$

Herbrandmodelle über der gewählten Signatur sind unendlich

Vollständigkeit

Lemma

Sei $\Gamma \Rightarrow \Delta$ nicht ableitbar. Dann gibt es Mengen $\Gamma^* \supseteq \Gamma$ und $\Delta^* \supseteq \Delta$ so dass

1. $\Gamma^* \cap \Delta^* = \emptyset$;
2. Wenn $\neg\varphi \in \Gamma^*$, dann $\varphi \in \Delta^*$;
   wenn $\neg\varphi \in \Delta^*$, dann $\varphi \in \Gamma^*$;
3. Wenn $\varphi \wedge \psi \in \Gamma^*$, dann $\varphi, \psi \in \Gamma^*$;
   wenn $\varphi \wedge \psi \in \Delta^*$, dann $\varphi \in \Delta^*$ oder $\psi \in \Delta^*$
4. Wenn $\exists x.\varphi(x) \in \Gamma^*$, dann gibt es Grundterm t mit $\varphi[t] \in \Gamma^*$
   wenn $\exists x.\varphi(x) \in \Delta^*$, dann $\varphi[t] \in \Delta^*$ für alle Grundterme t

Theorem (Vollständigkeit SK)

Wenn $\models \Gamma \Rightarrow \Delta$, dann $\vdash \Gamma \Rightarrow \Delta$ (jede gültige Sequenz ist ableitbar).

Sequenzenkalkül

Eine der wichtigsten Anwendungen des SK ist der Beweis der Kompaktheit von FO

Dies erfordert den Umgang mit unendlichen Satzmengen Π

Für eine Menge von Sätzen $\Pi \subseteq$FO erhält man die Π-*Erweiterung* des SK durch Hinzufügen der Regel

$$(\Pi) \quad \frac{\Gamma, \varphi \Rightarrow \Delta}{\Gamma \Rightarrow \Delta} \quad \varphi \in \Pi$$

Durch Anpassung der ursprünglichen Beweise zeigt man leicht:

Theorem (Korrektheit+Vollständigkeit erweiterter SK)

$\Gamma \Rightarrow \Delta$ in der Π-Erweiterung des SK ableitbar gdw. $\Pi \models \bigwedge \Gamma \Rightarrow \bigvee \Delta$

Sequenzenkalkül

Anmerkung:

- Es reichen wenige Regeln, um das vorgestellte SK-Kalkül auf FO mit Gleichheitsprädikat zu erweitern

- Der Vollständigkeitsbeweis wird dann etwas komplexer: man benötigt zusätzlich eine *Quotientenkonstruktion*

Kapitel 3.2: Rekursive Aufzählbarkeit, Kompaktheit und Löwenheim-Skolem

Rekursive Aufzählbarkeit

Theorem (Rekursive Aufzählbarkeit)

Für jede aufzählbare Signatur τ sind rekursiv aufzählbar:

- die Menge aller Tautologien aus $\text{FO}(\tau)$
- die Menge aller unerfüllbaren Formeln aus $\text{FO}(\tau)$

Beweis:

- die Menge aller τ-Formeln ist rekursiv aufzählbar,
  also auch die Menge aller SK-Beweise
- FO Formel $\varphi(x_1, \ldots, x_n)$ ist
  - Tautologie gdw. es SK-Beweis für $\emptyset \Rightarrow \forall x_1, \ldots, x_n.\varphi$ gibt,
  - unerfüllbar gdw. es SK-Beweis für $\exists x_1, \ldots, x_n.\varphi \Rightarrow \emptyset$ gibt

  (Korrektheit und Vollständigkeit des SK)

Rekursive Aufzählbarkeit

Korollar

Wenn τ mind. ein binäres Relationssymbol enthält, ist die Menge der erfüllbaren $\mathrm{FO}(\tau)$-Formeln nicht rekursiv aufzählbar.

Denn: Wären die erfüllbaren Formeln rekursiv aufzählbar, so wäre Erfüllbarkeit entscheidbar:

Um Erfüllbarkeit von φ zu prüfen, zähle simultan die erfüllbaren Formeln und die unerfüllbaren Formeln auf:

erfüllbar	unerfüllbar
φ_1	ψ_1
φ_2	ψ_2
$\vdots$	$\vdots$

Nach endlicher Zeit findet man Eingabeformel φ.

Theorembeweiser

Rekursive Aufzählbarkeit liefert *Semi-Entscheidbarkeit* für Gültigkeit (und Unerfüllbarkeit):

- wenn Eingabe Tautologie, dann terminiert der Algorithmus nach endlicher Zeit und antwortet “gültig”;
- wenn Eingabe keine Tautologie, dann keine Terminierung.

Auf diesem Prinzip beruhen moderne Theorembeweiser wie Vampire, Paradox, Spass; allerdings wird...

- meist Resolution verwendet (mit aufwendigen Optimierungstechniken)
- durch zusätzliche Verfahren in “vielen Fällen” auch Terminierung auf nicht-Tautologien erreicht

Theorembeweiser

Beachte:

wenn eine FO-Theorie Γ eine endliche Axiomatisierung Π hat, dann kann ein Theorembeweiser auch für Γ verwendet werden:

$$\varphi \in \Gamma \text{ gdw. } \bigwedge \Pi \rightarrow \varphi \text{ Tautologie}$$

Auch auf unendliche Axiomatisierungen können viele Beweiser angepasst werden

Man kann sie aber nicht verwenden, um Goldbachs Vermutung (oder andere zahlentheoretische Resultate) zu beweisen, denn

$$\mathsf{Th}(\mathbb{N}, +, *, 0, 1)$$

ist ja nicht axiomatisierbar.

Rekursive Aufzählbarkeit

Über endlichen Strukturen kehrt sich die Situation um:

Theorem (Rekursive Aufzählbarkeit, endliche Modelle)

Über endlichen Modellen gilt:

1. die Menge der erfüllbaren Formeln ist rekursiv aufzählbar, für jede aufzählbare Signatur τ
2. die Menge der unerfüllbaren Formeln ist nicht rekursiv aufzählbar, ebensowenig die Menge der Tautologien

Beweis in der Übungsgruppe.

Kompaktheit

Theorem (Kompaktheitssatz)

Für alle Mengen von Sätzen $\Gamma \subseteq \mathrm{FO}$ und Sätze $\varphi \in \mathrm{FO}$ gilt:

1. Γ ist erfüllbar gdw. jede endliche Teilmenge von Γ erfüllbar ist
2. $\Gamma \models \varphi$ gdw. endliches $\Delta \subseteq \Gamma$ existiert mit $\Delta \models \varphi$

Beweis einfach mittels Γ-Erweiterung des Sequenzenkalküls, also:

$$\Pi \Rightarrow \Delta \text{ ableitbar aus } \Gamma \quad \text{gdw.} \quad \Gamma \models \bigwedge \Pi \rightarrow \bigvee \Delta$$

Beachte: es wird hier eine syntaktische Eigenschaft (Kalkül!) in eine rein semantische (Erfüllbarkeit, Konsequenz!) übertragen.

Kompaktheit

Wir nutzen die Kompaktheit zum Beweis einiger wichtiger modell-theoretischer Resultate

Diese beziehen sich einerseits auf die Größe von Modellen:

- wie groß können die Modelle einer gegebenen Formel werden?
- gibt es Formeln, die nur in endlichen/unendlichen/abzählbaren/überabzählbaren Modellen erfüllbar sind?

Andererseits erlauben sie uns erste Beobachtungen bezüglich der Grenzen der Ausdrucksstärke von FO:

- kann ich eine Eigenschaft wie “das Modell ist endlich/unendlich/abzählbar/überabzählbar” in FO ausdrücken?

Unendliche Modelle

Theorem (unbeschränkte endliche Modelle)

Wenn ein FO-Satz φ beliebig große endliche Modelle besitzt (d.h. für jedes $n \geq 0$ gibt es Modell $\mathfrak{A}$ mit $|A| \geq n$), dann hat φ auch ein unendliches Modell.

Dieses Theorem impliziert eine Beschränkung der Ausdrucksstärke von FO:

> Es gibt keinen FO-Satz φ so dass $\mathfrak{A} \models \varphi$ gdw. $|A|$ endlich.
>
> "Endlichkeit ist nicht FO-ausdrückbar"

Für ein festes n ist "Modellgröße $\leq n$" aber natürlich leicht ausdrückbar:

$$\forall x_0, \ldots, x_n . \big(\bigvee_{0 \leq i < j \leq n} x_i = x_j \big)$$

Löwenheim-Skolem

Theorem (Aufsteigender Satz von Löwenheim-Skolem)

Wenn ein FO-Satz φ ein unendliches Modell besitzt, dann gibt es für jede Menge U ein Modell $\mathfrak{A}$ von φ mit $|A| \geq |U|$.

Beachte: Die Kardinalität von U ist beliebig!

Es folgt also z.B.:

wenn Γ unendliches Modell hat, dann auch überabzählbares Modell

(also ist auch Abzählbarkeit nicht FO-ausdrückbar)

Korollar (Nicht-Standardmodell der Arithmetik)

Th$(\mathbb{N}, +, \cdot, 0, 1)$ hat Modelle, die nicht isomorph zu $(\mathbb{N}, +, \cdot, 0, 1)$ sind.

Man kann sogar zeigen:

die Arithmetik $(\mathbb{N}, +, *, 0, 1)$ hat abzählbare Nichtstandardmodelle

Löwenheim-Skolem

Theorem (Absteigender Satz von Löwenheim-Skolem)

Wenn ein FO-Satz φ ein Modell besitzt und τ abzählbar ist,
dann hat φ auch ein endliches oder abzählbar unendliches Modell.

Es gibt also keine FO-Formeln, die nur überabzählbare Modelle haben.

Es folgt: Überabzählbarkeit nicht FO-ausdrückbar

Es folgt auch, dass es ein abzählbares Nichtstandardmodell
für die Arithmetik der reellen Zahlen $\mathsf{Th}(\mathbb{R}, +, *, 0, 1)$ gibt

Mehr zur Prädikatenlogik

Kapitel 3.3: Ausdrucksstärke /
Grundlagen von Ehrenfeucht-Fraïsse Spielen

Eigenschaften / Ausdrückbarkeit

In der Informatik ist die Analyse der Ausdrucksstärke von FO und anderen Logiken ein sehr wichtiges Thema, z.B.:

- Zusammenhang “SQL als FO”:

  Kann jede gewünschte Anfrage in SQL/FO ausgedrückt werden?

- FO in der Verifikation von Soft-/Hardware:

  Welche Systemeigenschaften können in FO beschrieben werden?

- Später: FO zur Definition von formalen Sprachen

  Welche formalen Sprachen können in FO definiert werden?

etc.

Eigenschaften / Ausdrückbarkeit

Statt Anfragen/Systemeigenschaften/Sprachen betrachten wir verallgemeinernd *Eigenschaften* von Strukturen

Sei R binäres Relationssymbol, T ternäres Relationssymbol

Beispiel 1: die Eigenschaft "$R^{\mathfrak{A}}$ ist eine Äquivalenzrelation"

ist FO-ausdrückbar:

$$\varphi = \forall x.R(x,x) \land \forall x,y.(R(x,y) \rightarrow R(y,x)) \land$$
$$\forall x,y,z.(R(x,y) \land R(y,z) \rightarrow R(x,z))$$

Beispiel 2: ebenso die Eigenschaft

"In $T^{\mathfrak{A}}$ sind die ersten beiden Spalten ein Primärschlüssel":

$$\varphi = \forall x,y,z,z'.\big(\,(\,T(x,y,z) \land T(x,y,z')\,)\,\rightarrow z = z'\,\big)$$

Eigenschaften / Ausdrückbarkeit

Definition Eigenschaft, Ausdrückbarkeit

Sei τ eine Signatur. Eine τ-*Eigenschaft* ist eine Klasse von τ-Strukturen, die unter Isomorphie abgeschlossen ist.
Die Eigenschaft P ist *FO-ausdrückbar* wenn es Satz $\varphi \in \text{FO}(\tau)$ gibt so dass $\mathfrak{A} \in P$ gdw. $\mathfrak{A} \models \varphi$ für alle Strukturen $\mathfrak{A}$.

Beispiele:

$$P_1 = \{\mathfrak{A} \mid R^{\mathfrak{A}} \text{ ist eine Äquivalenzrelation}\}$$

$$P_2 = \{\mathfrak{A} \mid \text{In } T^{\mathfrak{A}} \text{ sind die ersten beiden Spalten ein Primärschlüssel}\}$$

Eigenschaften, die nicht unter Isomorphie abgeschlossen sind,

- sind trivialerweise nicht FO-ausdrückbar
- "passen nicht zur Philosophie von FO".

Eigenschaften / Ausdrückbarkeit

Die Sätze von Löwenheim/Skolem und verwandte Resultate haben als Nebenprodukt Beschränkungen der Ausdrucksstärke von FO aufgezeigt

Nicht ausdrückbar sind folgende Eigenschaften:

- Endlichkeit von Strukturen
- Abzählbarkeit / Überabzählbarkeit von Strukturen

In der Informatik sind aber meist andere Eigenschaften relevant

Im folgenden: Werkzeuge zur Analyse der Ausdrucksstärke, wobei

- Ausdrückbarkeit meist leicht zu zeigen, Nicht-Ausdrückbarkeit schwierig!
- Man braucht unterschiedliche Methoden für endliche Modelle und beliebige Modelle

Zusammenhang

Zur Erinnerung:

ein ungerichteter Graph $G = (V, E)$ ist *zusammenhängend* wenn es für alle Knoten $v, v' \in V$ eine Knotenfolge $v_1, \ldots, v_n$ gibt so dass $v = v_1$, $v_n = v'$ und $\{v_i, v_{i+1}\} \in E$ für $1 \leq i < n$

Ungerichtete Graphen sind nichts anderes als $\{E\}$-Stukturen, E binäres Relationssymbol, das symmetrisch interpretiert wird.

In der Mathematik wird Nicht-Ausdrückbarkeit oft über Kompaktheit bewiesen

Theorem

Zusammenhang von ungerichteten Graphen ist nicht FO-ausdrückbar.

Zusammenhang

Kann man also Zusammenhang auch in SQL nicht ausdrücken?

Leider können wir das nicht aus dem vorigen Resultat folgern, denn

- Datenbankinstanzen entsprechen *endlichen* Modellen
- der Kompaktheitssatz gilt auf endlichen Modellen nicht: für

$$\Gamma = \{\exists x_1, \ldots, x_n. \bigwedge_{i \neq j} x_i \neq x_j \mid n \geq 2\}$$

(“es gibt unbeschränkt viele Elemente”) gilt:

- jede endliche Teilmenge von Γ ist in endlichem Modell erfüllbar
- Γ hat kein endliches Modell

Wir brauchen ein besseres Werkzeug zur Analyse der Ausdrucksstärke!

Ehrenfeucht-Fraïsse Spiele

Ehrenfeucht-Fraïsse Spiele sind eine elegante Beweistechnik, die es erlaubt, die *Nicht*-Ausdrückbarkeit von Eigenschaften in FO (und anderen Logiken) nachzuweisen.

Eine für die Informatik besonders wichtige Eigenschaft:

> Ehrenfeucht-Fraïsse Spiele funktionieren auf endlichen und unendlichen Modellen gleichermaßen

Wie wir gesehen haben, gilt das für viele andere Resultate nicht (z.B. Kompaktheit, rekursive Aufzählbarkeit von Tautologien)

Ehrenfeucht-Fraïsse Spiele

- Zwei Spieler: Spoiler und Duplikator
- Das Spielbrett besteht aus zwei τ-Strukturen $\mathfrak{A}$ und $\mathfrak{B}$ (endlich oder unendlich)
- Die Spieler wechseln sich ab, Spoiler beginnt
- Die zu spielende Rundenzahl k ist beliebig, aber vorher festgelegt
- In jeder Runde wählt Spoiler zunächst eine Stuktur ($\mathfrak{A}$ oder $\mathfrak{B}$), dann ein Element der gewählten Struktur

  Duplikator antwortet mit einem Element der anderen Struktur
- Im Prinzip: Spoiler möchte zeigen, dass $\mathfrak{A}$ und $\mathfrak{B}$ unterschiedlich sind, Duplikator dass sie gleich sein
- Die genaue Gewinnbedingung werden wir gleich definieren.

Gewinnbedingung

Der Einfachheit halber arbeiten wir im folgenden mit relationalen Signaturen

Wenn $\mathfrak{A}$ Struktur und $S \subseteq A$, so ist $\mathfrak{A}|_S$ die *Einschränkung* von $\mathfrak{A}$ auf S:

- das Universum von $\mathfrak{A}|_S$ ist S
- für alle n-stelligen Relationssymbole R:

$$R^{\mathfrak{A}|_S} = \{(a_1, \dots, a_n) \in R^{\mathfrak{A}} \mid a_1, \dots, a_n \in S\}$$

Definition Partieller Isomorphismus

Seien $\mathfrak{A}$ und $\mathfrak{B}$ τ-Strukturen und $\delta : A \to B$ eine partielle Funktion mit Definitionsbereich $\mathsf{dom}(\delta)$ und Wertebereich $\mathsf{ran}(\delta)$. Dann ist δ ein *partieller Isomorphismus* wenn δ ein Isomorphismus von $\mathfrak{A}|_{\mathsf{dom}(\delta)}$ nach $\mathfrak{B}|_{\mathsf{ran}(\delta)}$ ist.

Gewinnbedingung

Gewinner eines EF-Spieles:

- Angenommen, es wurden alle k Runden gespielt und in Runde i wurden die Elemente $a_i \in A$ und $b_i \in B$ ausgewählt
- Wenn die erreichte Menge

$$\{(a_1, b_1), \ldots, (a_k, b_k)\}$$

  ein partieller Isomorphismus ist, gewinnt Duplikator.
- Sonst gewinnt Spoiler.

Uns interessiert weniger der Gewinner eines einzelnen Spielverlaufs, sondern hauptsächlich der Gewinner bei optimaler Spielweise

Gewinnstrategien

- Das Spiel auf $\mathfrak{A}$,$\mathfrak{B}$ mit k-Zügen bezeichnen wir mit $\mathcal{G}_k(\mathfrak{A}, \mathfrak{B})$
- Ein Spieler *hat eine Gewinnstrategie* für $\mathcal{G}_k(\mathfrak{A}, \mathfrak{B})$ wenn er dieses Spiel gewinnen kann, egal was der andere Spieler tut
- Gewinnstrategien für $\mathcal{G}_k(\mathfrak{A}, \mathfrak{B})$ kann man anschaulich als endliche Spielbäume der Tiefe k darstellen
- Für jedes Spiel $\mathcal{G}_k(\mathfrak{A}, \mathfrak{B})$ hat Spoiler oder Duplikator eine Gewinnstrategie (denn das gilt für alle endlichen 2-Personen-Spiele, in denen kein Unentschieden möglich ist)

Beispiele

Gewinnstrategien

Beachte:

- Abwechselnde Züge entsprechen Quantorenalternierungen
- Gewinnstrategien für Spoiler und Duplikator sind dual

Gewinnstrategie Spoiler:

$\exists$ Zug Spoiler so dass
  $\forall$ Züge Duplikator gilt
    $\exists$ Zug Spoiler so dass
      $\ddots$
        $\forall$ Züge Duplikator gilt
          Spiel ist kein part. Isom.

Gewinnstrategie Duplikator:

$\forall$ Züge Spoiler gilt
  $\exists$ Zug Duplikator so dass
    $\forall$ Züge Spoiler gilt
      $\ddots$
        $\exists$ Zug Duplikator so dass
          Spiel ist part. Isom.

Quantorenrang

Wir stellen nun den Zusammenhang zwischen EF und FO her

Die Anzahl der Spielrunden entspricht dabei dem Quantorenrank

Definition Quantorenrang

Der *Quantorenrang* $\mathsf{qr}(\varphi)$ einer Formel φ ist die Schachtelungstiefe von Quantoren in φ. Formal:

- wenn φ ein Atom, dann $\mathsf{qr}(\varphi) = 0$
- $\mathsf{qr}(\neg\varphi) = \mathsf{qr}(\varphi)$
- $\mathsf{qr}(\varphi \wedge \psi) = \mathsf{qr}(\varphi \vee \psi) = \max\{\mathsf{qr}(\varphi), \mathsf{qr}(\psi)\}$
- $\mathsf{qr}(\exists x.\varphi) = \mathsf{qr}(\forall x.\varphi) = \mathsf{qr}(\varphi) + 1$

Beispiel:

$$\mathsf{qr}(\exists x.(\forall y.P(x,y) \vee \exists z.\forall y.Q(x,y,z))) = 3$$

Quantorenrang

Folgende Beobachtung über den Quantorenrang werden wir benötigen:

Lemma über paarweise Äquivalenz

Sei τ eine endliche Signatur, m eine Stelligkeit und k ein Quantorenrang. Es gibt nur endlich viele Formeln $\varphi \in \mathsf{FO}(\tau)$ mit m freien Variablen und $\mathsf{qr}(\varphi) = k$, die paarweise nicht äquivalent sind.

Beispiel: $\tau = \{P\}$, P 1-stellig

Für $k = 0$ gibt es vier Äquivalenzklassen:

$$P(x), \quad \neg P(x), \quad P(x) \wedge \neg P(x), \quad P(x) \vee \neg P(x)$$

Es gilt z.B. $P(x) \vee P(x) \equiv P(x)$

Für $k = 1$ gibt es schon 16 Äquivalenzklassen (Übung!)

Ehrenfeucht-Fraïsse Theorem

Theorem (Ehrenfeucht-Fraïsse)

Seien $\mathfrak{A}$ und $\mathfrak{B}$ τ-Strukturen. Für alle $k \geq 0$ sind folgende Aussagen äquivalent:

1. $\mathfrak{A} \models \varphi$ gdw. $\mathfrak{B} \models \varphi$ für alle Sätze $\varphi \in \text{FO}(\tau)$ mit $\text{qr}(\varphi) \leq k$
2. Duplikator hat Gewinnstrategie für $\mathcal{G}_k(\mathfrak{A}, \mathfrak{B})$.

Um einen Induktionsbeweis durchführen zu können, müssen wir Spiele betrachten, die bereits einige Runden gespielt wurden:

Angenommen, nach i Zügen ist *Spielposition* $\{(a_1, b_1), \ldots, (a_i, b_i)\}$ erreicht

Das verbleibende Restspiel mit k Zügen bezeichnen wir mit

$$\mathcal{G}_k(\mathfrak{A}, a_1, \ldots, a_i, \mathfrak{B}, b_1, \ldots, b_i)$$

Gewinnstrategien für Teilspiele genauso definiert wie für vollständige Spiele

Ehrenfeucht-Fraïsse Theorem

Wir beweisen nun folgende allgemeinere Aussage per Induktion über k

Theorem

Seien $\mathfrak{A}$ und $\mathfrak{B}$ τ-Strukturen, $\overline{a} = a_1, \ldots, a_r \in A$ und $\overline{b} = b_1, \ldots, b_r \in B$.
Für alle $k \geq 0$ sind folgende Aussagen äquivalent:

1. $\mathfrak{A} \models \varphi[\overline{a}]$ und $\mathfrak{B} \not\models \varphi[\overline{b}]$ für eine Formel $\varphi(\overline{x}) \in \mathsf{FO}(\tau)$ mit $\mathsf{qr}(\varphi) \leq k$
2. Spoiler hat Gewinnstrategie für $\mathcal{G}_k(\mathfrak{A}, \overline{a}, \mathfrak{B}, \overline{b})$.

Unterschiede:

1. bereits begonnene Spiele, also Formeln mit freien Variablen

2. Gewinnstrategien für Spoiler und Unterscheidbarkeit durch Formeln
statt Duplikatior und Ununterscheidbarkeit

Offensichtlich folgt das Theorem von Ehrenfeucht-Fraïsse