Automatisches Beweisen—Vertiefung
First Order Logic

Christoph Zengler

Arbeitsbereich Symbolisches Rechnen
Prof. Dr. Wolfgang Küchlin
Universität Tübingen

30. Oktober 2012

Logischer Formalismus

Syntax

Wie werden Formeln gebildet?

- klassisch-mathematisch: bestimmte ausgezeichnete Zeichenreihen
- informatisch: Sprache einer Grammatik

Semantik

Was ist die Bedeutung einer Formel?

- Allgemein: Abbildung in einen (bekannten) Semantik-Bereich

Kalkül

Wie kann der Wahrheitswert einer Formel bestimmt werden?

- Inferenzregeln, Algorithmen

Aussagenlogik vs. Logik erster Stufe

Aussagenlogik

- Nur atomare Aussagen, die wahr oder falsch sind, z.B.
  - *“5 ist eine Primzahl”*
  - *“7 ist gerade”*
  - *“Thomas ist ein Student”*
- Variablen können nicht quantifiziert werden

Logik erster Stufe

- Atome sind Prädikate, die aus Termen und Variablen über einem Universum bestehen, z.B.
  - $x < y$
  - $\texttt{istStudent}(x)$
  - $f(x, y) \geq g(x) + g(y)$
- Variablen können mit $\exists$ und $\forall$ quantifiziert werden

Syntax der First Order Logic (FOL)

EBNF-Grammatik für Formeln erster Stufe

Term	=	*Termvariable*	$\in \mathcal{V}_T$
	|	*Funktionssymbol*({Term})	Funktionsapplikation

Formel	=	$\top$ | $\bot$	Konstanten
	|	$\neg$ Formel	Negation
	|	*Prädikatssymbol*({Term})	Prädikatapplikation
	|	Formel $\wedge$ Formel	Konjunktion
	|	Formel $\vee$ Formel	Disjunktion
	|	Formel $\rightarrow$ Formel	Implikation
	|	Formel $\leftrightarrow$ Formel	Äquivalenz
	|	$\forall$ Termvariable [Formel]	Allquantor
	|	$\exists$ Termvariable [Formel]	Existenzquantor
	|	(Formel)	Klammerung

Anmerkungen zur Syntax

- Unter Atomaren Formeln fassen wir Prädikate und Konstanten zusammen
- Jedes Funktionssymbol und jedes Prädikatssymbol hat eine Stelligkeit (Arität), die angibt wie viele Terme als Parameter übergeben werden
  - Notation $\mathrm{f}^{(2)}$: Funktionssymbol f hat Stelligkeit 2
- Funktionen mit Stelligkeit 0 nennen wir auch Konstanten, Prädikate mit Stelligkeit 0 nennen wir Boolesche Konstanten
- Formeln repräsentieren Wahrheitswerte, Terme repräsentieren Objekte (z.B. Zahlen, Studenten, . . .)
- Relationen werden auch Prädikate genannt
- Logik erster Stufe wird auch Prädikatenlogik genannt
- In FOL dürfen wir nur über Termvariablen quantifizieren, will man z.B. über Boolesche Konstanten, Funktionen oder Prädikate quantifizieren, muss man eine Logik höherer Ordnung (HOL) benutzen.

First Order Logic—Beispiele

Beispiel (Syntax der FOL)

Sprache: $(+^{(2)}, -^{(1)}, 0^{(0)}, \leq^{(2)}, =^{(2)})$

Wohlgeformte Terme:

- $x + 0$, $-x$, -0, $x + y$

Wohlgeformte Atomare Formeln

- $x + 0 \leq -x$
- $x + 0 = 0$
- $x + -y \leq -z$

Wohlgeformte Formeln

- $x + 0 \leq -x \wedge x + 0 = 0$
- $\exists z[x + -y \leq -z]$

Quantoren

- Quantoren $\exists x$ und $\forall x$ binden die Variable x (lokale Definition)
- Eine Variable kann gleichzeitig gebunden und frei in einer Formel vorkommen

Definition (Freie und gebundene Variablen)

- Menge der freien Variablen einer Formel φ: $\mathtt{free}(\varphi)$
- Menge der gebundenen Variablen einer Formel φ: $\mathtt{bound}(\varphi)$

Beispiel (Freie und gebundene Variablen)

$$\varphi = \exists a[a = 12 \wedge b > a \wedge \forall u \exists b[b + u = 24 \wedge x = y]]$$

- $\mathtt{free}(\varphi) = \{b, x, y\}$
- $\mathtt{bound}(\varphi) = \{a, b, u\}$

- Wir erlauben die abkürzende Schreibweise $Q(x_1, \ldots, x_n)[\varphi]$ für $Qx_1[\ldots[Qx_n[\varphi]]]$ mit $Q \in \{\exists, \forall\}$

Semantik der FOL

1. Belegung $\beta_D : \mathcal{V}_T \to D$ der Termvariablen
2. Interpretation $\mathcal{M}$ bestehend aus 3 Teilen:

Definition (Das Universum D)

nicht-leere Menge D (domain) der Individuen, über die man spricht. D.h. alle Terme evaluieren zu einem Wert in D.

Definition (Interpretation der Funktionssymbole)

Jedem n-stelligen Funktionssymbol f wird eine Funktion $f_{\mathcal{M}} : D^n \to D$ zugeordnet

Definition (Interpretation der Prädikatssymbole)

Jedem n-stelligen Prädikatssymbol P wird eine Boolesche Funktion $P_{\mathcal{M}} : D^n \to \{\texttt{true}, \texttt{false}\}$ zugeordnet, d.h. $P_{\mathcal{M}} \subseteq D^n$

Evaluation von Termen

Algorithmus: $\texttt{termval}(\mathcal{M}, \beta_D, t)$

Eingabe: Interpretation $\mathcal{M}$, Belegung β_D, Term t
Ausgabe: Evaluation von t unter D und β_D

$$\texttt{termval}(\mathcal{M}, \beta_D, t) = t \texttt{ match}$$
$$\quad x \in \mathcal{V}_T \rightsquigarrow \beta_D(x)$$
$$\mid f(t_1, \ldots, t_n) \rightsquigarrow f_{\mathcal{M}}(\texttt{termval}(\mathcal{M}, \beta_D, t_1), \ldots, \texttt{termval}(\mathcal{M}, \beta_D, t_n))$$

Beispiel (Evaluation von Termen)

- Funktionssymbole: $\diamond^{(0)}, \ominus^{(1)}, \oplus^{(2)}$
- Interpretation $\mathcal{M}$ mit Universum $\mathbb{Z}$ und Interpretationen
  - $\diamond_{\mathcal{M}} = 0$
  - $\ominus_{\mathcal{M}}(x) = -x$
  - $\oplus_{\mathcal{M}}(x, y) = x + y$
- $\beta_D = \{x \mapsto 2, y \mapsto -4\}$

$\texttt{termval}(\oplus(x, \ominus(\oplus(y, \diamond)))) = 6 \in \mathbb{Z}$

Evaluation von FOL Formeln

Algorithmus: $\texttt{holds}(\mathcal{M}, \beta_D, \psi)$

Eingabe: Interpretation $\mathcal{M}$, Belegung β_D der Termvariablen, Formel ψ
Ausgabe: Evaluation von ψ unter $\mathcal{M}$ und β_D

$\texttt{holds}(\mathcal{M}, \beta_D, \psi) = \psi\ \texttt{match}$

$\top \rightsquigarrow \texttt{true}$

$\mid \bot \rightsquigarrow \texttt{false}$

$\mid P(t_1, \ldots, t_n) \rightsquigarrow P_{\mathcal{M}}(\texttt{termval}(\mathcal{M}, \beta_D, t_1), \ldots, \texttt{termval}(\mathcal{M}, \beta_D, t_n))$

$\mid \neg\varphi \rightsquigarrow \texttt{if holds}(\mathcal{M}, \beta_D, \varphi)\ \texttt{then false else true}$

$\mid \varphi_1 \wedge \varphi_2 \rightsquigarrow \texttt{if holds}(\mathcal{M}, \beta_D, \varphi_1)\ \texttt{and holds}(\mathcal{M}, \beta_D, \varphi_2)\ \texttt{then true else false}$

$\mid \varphi_1 \vee \varphi_2 \rightsquigarrow \texttt{holds}(\mathcal{M}, \beta_D, \neg(\neg\varphi_1 \wedge \neg\varphi_2))$

$\mid \varphi_1 \rightarrow \varphi_2 \rightsquigarrow \texttt{holds}(\mathcal{M}, \beta_D, \neg\varphi_1 \vee \varphi_2)$

$\mid \varphi_1 \leftrightarrow \varphi_2 \rightsquigarrow \texttt{holds}(\mathcal{M}, \beta_D, \varphi_1 \rightarrow \varphi_2 \wedge \varphi_2 \rightarrow \varphi_1)$

$\mid \forall x[\varphi] \rightsquigarrow \texttt{for all } a \in D\texttt{: holds}(\mathcal{M}, \beta_D \cup \{x \mapsto a\}, \varphi)$

$\mid \exists x[\varphi] \rightsquigarrow \texttt{exists } a \in D\texttt{: holds}(\mathcal{M}, \beta_D \cup \{x \mapsto a\}, \varphi)$

Evaluation von FOL Formeln—Beispiel

Beispiel (Evaluation von FOL Formeln)

- Funktionssymbole: $\diamond^{(0)}, \ominus^{(1)}, \oplus^{(2)}$
- Prädikatsssymbole: $\sqsubseteq^{(2)}, \doteq^{(2)}$
- Interpretation $\mathcal{M}$ mit Universum $\mathbb{Z}$ und Interpretationen
  - $\diamond_{\mathcal{M}} = 0$
  - $\ominus_{\mathcal{M}}(x) = -x$
  - $\oplus_{\mathcal{M}}(x, y) = x + y$
  - $\sqsubseteq_{\mathcal{M}}(x, y) = x \leq y$
  - $\doteq_{\mathcal{M}}(x, y) = x \neq y$
- $\beta_D = \{x \mapsto 2, y \mapsto -4\}$
- $\psi = \sqsubseteq(\oplus(x, \ominus(\oplus(y, \diamond))), \ominus(x)) \vee \neg(\doteq(\oplus(x, y), \diamond))$

$$
\begin{aligned}
\texttt{holds}(\mathcal{M}, \beta_D, \psi) &= (x + -(y + 0) \leq -x) \vee \neg(x + y \neq 0) \\
&= (2 + -(-4 + 0) \leq -2) \vee \neg(2 + -4 \neq 0) \\
&= 6 \leq -2 \vee -2 = 0 \\
&= \texttt{false}
\end{aligned}
$$

Freie Variablen

Theorem (Formeln unter gleicher Belegung der freien Variablen)

Gilt für eine Formel φ

- *für alle* $x \in \texttt{free}(\varphi)$, *dass* $\beta_D(x) = \beta'_D(x)$

so gilt auch

- $\texttt{holds}(\mathcal{M}, \beta_D, \varphi) = \texttt{holds}(\mathcal{M}, \beta'_D, \varphi)$ *für beliebige Interpretationen* $\mathcal{M}$.

- Ein Term oder eine Formel ohne Variablen heißt variablenfrei oder ground.
- Eine Formel ohne freie Variablen heißt Satz.

Theorem (Sätze unter beliebigen Belegungen)

Ist φ *ein Satz, so gilt für beliebige Belegungen* β_D *und* β'_D, *dass* $\texttt{holds}(\mathcal{M}, \beta_D, \varphi) = \texttt{holds}(\mathcal{M}, \beta'_D, \varphi)$.

Gültigkeit

Definition (Gültigkeit)

Eine FOL Formel φ heißt gültig oder valide, wenn für alle möglichen Interpretationen $\mathcal{M}$ und Belegungen β_D gilt, dass $\texttt{holds}(\mathcal{M}, \beta_D, \varphi) = \texttt{true}$.

Beispiel (Gültige Formel)

$$\forall x[P(x)] \rightarrow P(a)$$

(gilt ein Prädikat P für alle Objekte, so gilt es auch für ein spezielles Objekt a)

Vorsicht!

Der Quantor $\forall x$ und sein Scope sind wichtig. Weder

- $P(x) \rightarrow P(a)$ noch
- $\forall x[P(x) \rightarrow P(a)]$

sind gültig.

Erfüllbarkeit

Definition (Erfüllbarkeit)

Eine Interpretation $\mathcal{M}$ erfüllt eine FOL Formel φ (φ gilt in $\mathcal{M}$), wenn für *alle* Belegungen β_D gilt, dass $\texttt{holds}(\mathcal{M}, \beta_D, \varphi) = \texttt{true}$.

- $\mathcal{M}$ erfüllt eine Menge von FOL Formeln Γ (Γ gilt in $\mathcal{M}$), wenn $\mathcal{M}$ jede Formel $\varphi \in \Gamma$ erfüllt
- Eine FOL Formel φ ist erfüllbar, wenn eine Interpretation $\mathcal{M}$ *existiert*, die sie erfüllt.

Theorem (Gültigkeit von Sätzen)

Ein Satz φ ist gültig, genau dann wenn $\neg\varphi$ unerfüllbar ist.

⚠ Vorsicht!

Obiger Satz gilt nicht für Formeln mit freien Variablen:

$P(x) \vee \neg P(y)$ ist nicht gültig, $\neg P(x) \wedge P(y)$ ist jedoch unerfüllbar.

Modelle

Definition (Modell)

Eine Interpretation $\mathcal{M}$, die eine Menge Γ von Formeln erfüllt, heißt Modell von Γ.

Notationen:

- $\Gamma \models \varphi$: φ gilt in allen Modellen von Γ (d.h. Jedes Modell von Γ ist auch ein Modell von φ)
- Wir schreiben $\models \varphi$ statt $\varnothing \models \varphi$ (gilt in allen Modellen)
- $\models_{\mathcal{M}} \varphi$: $\mathcal{M}$ erfüllt φ

Konsequenz:

- Γ ist unerfüllbar, gdw. $\Gamma \models \bot$

⚠ Vorsicht!

Im Gegensatz zur Aussagenlogik, gilt in FOL nicht, dass $\{\varphi_1, \ldots, \varphi_n\} \models \varphi$ äquivalent ist zu $\models \varphi_1 \wedge \cdots \wedge \varphi_n \to \varphi$

$$\{P(x)\} \models P(y) \text{ aber nicht } \models P(x) \to P(y)$$

Abschlüsse

Oft is es angenehmer mit Sätzen anstelle von beliebigen Formeln zu arbeiten. Wenn z.B. alle Formeln φ_i Sätze sind, so gilt auch in FOL

$$\{\varphi_1, \ldots, \varphi_n\} \models \varphi \text{ ist äquivalent zu } \models \varphi_1 \wedge \cdots \wedge \varphi_n \rightarrow \varphi.$$

Definition (Universeller Abschluss)

Allquantifiziere alle freien Variablen einer Formel φ.

- Notation: $\underline{\forall}\varphi$

- φ ist gültig, gdw. $\underline{\forall}\varphi$ gültig ist.

Definition (Existentieller Abschluss)

Existenzquantifiziere alle freien Variablen einer Formel φ.

- Notation: $\underline{\exists}\varphi$

Substitution in Termen

Konvention: Substitution σ bildet immer alle freien Variablen einer Formel ab. Variablen, die nicht substitutiert werden, werden auf sich selber abgebildet.

Algorithmus: $\mathtt{tsubst}(\sigma, t)$

Eingabe: Abbildung σ von Variablen $\in \mathcal{V}_T$ zu Termen, Term t
Ausgabe: Term $t\sigma$ (Substitution auf t ausgeführt)

$\mathtt{tsubst}(\sigma, t) = t$ match

$x \in \mathcal{V}_T \rightsquigarrow \sigma(x)$

$\mid f(t_1, \ldots, t_n) \rightsquigarrow f(\mathtt{tsubst}(\sigma, t_1), \ldots, \mathtt{tsubst}(\sigma, t_n))$

Beispiel (Substitution in Termen)

- $t = (x + 24) \cdot (y - a)$
- $\sigma = \{x \mapsto 12, y \mapsto x + z\}$

$$\mathtt{tsubst}(\sigma, t) = (12 + 24) \cdot ((x + z) - a)$$

Substitution in Formeln

⚠ Nicht so einfach wie gedacht!

- Gebundene Variablen dürfen nicht substitutiert werden
- Substitution kann Variablen einführen, die bereits gebunden sind

Beispiel (Problemfälle)

- Substitution von x in $\forall x[x = x]$ darf keine Auswirkung haben
- Substitution $\{y \mapsto x\}$ in $\exists x[x + 1 = y]$ würde aus freier Variable y eine gebundene machen

Lösungsidee

- Gebundene Variablen sind nur Platzhalter

$\Rightarrow$ können umbenannt werden (Alpha Konversion)

Substitution in Formeln—Vorgehen

Algorithmus: $\mathtt{subst}(\sigma, \psi)$

Eingabe: Substitution σ, Formel ψ
Ausgabe: $\psi\sigma$

$\mathtt{subst}(\sigma, \psi) = \psi$ match

$\top \rightsquigarrow \top$

| $\bot \rightsquigarrow \bot$

| $P(t_1, \ldots, t_n) \rightsquigarrow P(\mathtt{tsubst}(\sigma, t_1), \ldots, \mathtt{tsubst}(\sigma, t_n))$

| $\neg\varphi \rightsquigarrow \neg\mathtt{subst}(\sigma, \varphi)$

| $\varphi_1 \wedge \varphi_2 \rightsquigarrow \mathtt{subst}(\sigma, \varphi_1) \wedge \mathtt{subst}(\sigma, \varphi_2)$

| $\varphi_1 \vee \varphi_2 \rightsquigarrow \mathtt{subst}(\sigma, \varphi_1) \vee \mathtt{subst}(\sigma, \varphi_2)$

| $\varphi_1 \to \varphi_2 \rightsquigarrow \mathtt{subst}(\sigma, \varphi_1) \to \mathtt{subst}(\sigma, \varphi_2)$

| $\varphi_1 \leftrightarrow \varphi_2 \rightsquigarrow \mathtt{subst}(\sigma, \varphi_1) \leftrightarrow \mathtt{subst}(\sigma, \varphi_2)$

| $\forall x[\varphi] \rightsquigarrow \mathtt{substq}(\sigma, \forall, x, \varphi)$

| $\exists x[\varphi] \rightsquigarrow \mathtt{substq}(\sigma, \exists, x, \varphi)$

$\mathtt{substq}(\sigma, q, v, \varphi) =$

$x = \mathtt{if}$ exists $y \neq v \in \mathtt{free}(\varphi)$ with $v \in \mathtt{free}(\sigma(y))$ then v' else v

return $qx[\mathtt{subst}(\sigma \cup \{v \mapsto x\}, \varphi)]$

Substitution in Formeln—Eigenschaften

⚠ Simultane Substitution

Substitution ist simultan, d.h. alle Substitutionen werden gleichzeitig ausgeführt.

Corollary

Wenn eine Formel φ gültig ist, so auch $\varphi\sigma$ für beliebige Substitutionen σ.

Notationen:

- Wir sagen statt Substitution auch oft Instanziierung
- Schreibweise: $\varphi\sigma$ statt $\texttt{subst}(\sigma, \varphi)$

Demo

```
val phi1 = "?[X]:X = X".fol
val phi2 = "?[X]: plus(X,1) = Y".fol

phi1.substitute(FOLVariable("X"), FOLVariable("Y"))
... = ∀(X)[X = X]

phi2.substitute(FOLVariable("Y"), FOLVariable("X"))
... = ∃(X_0)[plus(X_0,1) = X]
```

Negationsnormalform

Definition (NNF)

Eine FOL Formel φ ist in NNF, wenn

1. nur die Booleschen Operatoren $\neg, \wedge$ und $\vee$ in φ vorkommen
2. $\neg$ nur vor Atomaren Formeln steht

Beispiel (Negationsnormalform)

- $\neg(P(x) \vee (P(y) \wedge \neg R(z)))$ nicht in NNF
- $\neg P(x) \wedge (\neg P(y) \vee R(z))$ ist in NNF

Idee!

Benutze Dualität der Quantoren:

$$\forall x[\varphi] \equiv \neg\exists x[\neg\varphi]$$

$$\exists x[\varphi] \equiv \neg\forall x[\neg\varphi]$$

Negationsnormalform—Algorithmus

Algorithmus: $\mathtt{nnf}(\psi)$

Eingabe: FOL Formel ψ
Ausgabe: NNF von ψ

$\mathtt{nnf}(\psi) = \psi$ `match`

alle Regeln von `nnf` *Algorithmus der Aussagenlogik +*

- $\forall x[\varphi] \rightsquigarrow \forall x[\mathtt{nnf}(\varphi)]$
- $\exists x[\varphi] \rightsquigarrow \exists x[\mathtt{nnf}(\varphi)]$
- $\neg\forall x[\varphi] \rightsquigarrow \exists x[\mathtt{nnf}(\neg\varphi)]$
- $\neg\exists x[\varphi] \rightsquigarrow \forall x[\mathtt{nnf}(\neg\varphi)]$

Pränexnormalform (PNF)

Definition (PNF)

Eine FOL Formel φ in NNF ist in PNF, wenn

- sie von der Form $Q_1 x_1\ Q_2 x_2 \ldots Q_n x_n \psi$ ist, mit
- $Q_i \in \{\exists, \forall\}$ und
- ψ quantorenfrei

d.h. alle Quantoren kommen nur ganz außen vor

Notation:

- der quantorenfreie Anteil ψ einer PNF Formel φ heißt auch Matrix. $\psi = \mathrm{mat}(\varphi)$.

Beispiel (PNF)

- $\exists x[P(x)] \to \exists y[P(y) \wedge \forall z[P(z)]]$ nicht in PNF
- $\forall x \exists y \forall z[P(x) \wedge P(y) \vee P(z)]$ ist in PNF

Pränexnormalform—Algorithmus

Implementierung sehr technisch, hier nur die Skizze

1. Bringe Formel zuerst in NNF
2. Quantoren werden sukzessive nach außen gezogen
   - Umbenennungen können nötig sein:
     - $P(x) \wedge \exists x[R(x)] \not\equiv \exists x[P(x) \wedge R(x)]$
     - $P(x) \wedge \exists x[R(x)] \equiv \exists x'[P(x) \wedge R(x')]$
   - Quantoren können in speziellen Fällen reduziert werden (Mini Scoping):
     - $\forall x[\varphi_1] \wedge \forall y[\varphi_2] \equiv \forall z[\varphi_1[x/z] \wedge \varphi_2[y/z]]$
     - $\exists x[\varphi_1] \vee \exists y[\varphi_2] \equiv \exists z[\varphi_1[x/z] \vee \varphi_2[y/z]]$

Übung

Beweisen Sie die folgenden Äquivalenzen:

- $\forall x[\varphi_1] \wedge \forall y[\varphi_2] \equiv \forall z[\varphi_1[x/z] \wedge \varphi_2[y/z]]$
- $\exists x[\varphi_1] \vee \exists y[\varphi_2] \equiv \exists z[\varphi_1[x/z] \vee \varphi_2[y/z]]$

Pränexnormalform—Eigenschaften & Demo

- PNF Konversion erhält freie Variablen und deren Namen
- Bei einer Formel φ in PNF ist die Matrix $\mathrm{mat}(\varphi)$ im Scope eines jeden Quantors

$\Rightarrow$ Keine Variable kann gleichzeitig frei und gebunden vorkommen: $\mathtt{bound}(\varphi) \cap \mathtt{free}(\varphi) = \varnothing$

Demo

```
val phi = "p(X) & ?[X]: r(X)".fol

phi.pnf
```

```
... = ∃X0[(p(X) ∧ r(X0))]
```

Die Unentscheidbarkeit der FOL

- **Aussagenlogik:** Entscheidung über Wahrheitswert einer Formel mit Hilfe von Wahrheitstafeln $\Rightarrow$ **entscheidbar**
- **First Order Logic:** Im Allgemeinen unmöglich! Es gibt Formeln, die nur unendliche Modelle besitzen.

Beispiel (Unentscheidbarkeit von FOL)

$\varphi = \forall(u, v, w)[(P(u, v) \land P(v, w)) \rightarrow P(u, w)] \land \forall x[\neg P(x, x)] \land \forall y[P(y, f(y))]$

Unendliches Modell mit Universum $\mathbb{N}$

- $f_M = \mathbb{N} \rightarrow \mathbb{N} : n \mapsto n + 1$
- $P_M = \{(x, y) \in \mathbb{N}^2 \mid x < y\}$
- β_D beliebig, da alle Variablen gebunden

Endliches Modell ist nicht möglich!

Theorem (Satz von Church)

Es gibt keinen Algorithmus, der für eine beliebige gegebene FOL Formel φ in endlich vielen Schritten entscheidet, ob φ erfüllbar ist, oder nicht.

... aber nicht verzagen!

Wir werden im Folgenden sehen, dass die FOL immerhin semi-entscheidbar ist.

Was heißt das?

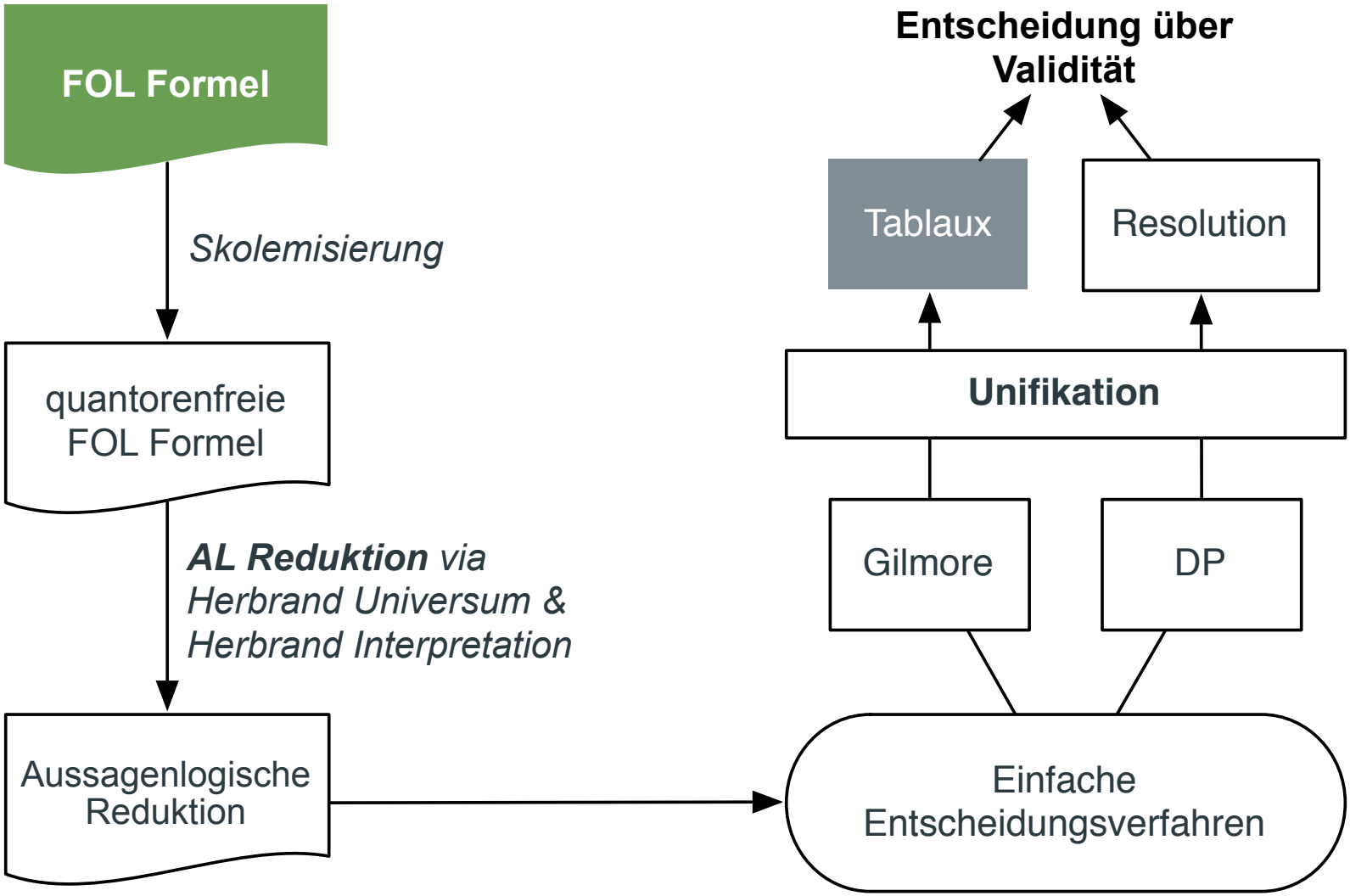
- Für eine unerfüllbare Formel können wir ein Verfahren angeben, das die Unerfüllbarkeit in endlicher Zeit nachweist.
- Problem: Wir können keine obere Grenze für die Laufzeit dieses Verfahrens angeben.

$\Rightarrow$ Wir wissen nicht, ob das Verfahren noch arbeitet, oder ob die Formel erfüllbar ist.

i Was können wir semi-entscheiden?

- Ist eine Formel eine Kontradiktion oder eine Tautologie (via Negation)
- Allgemeine Erfüllbarkeit kann nicht entschieden werden!

Der weitere Plan



Skolemisierung—Motivation

- PNF trennt Quantoren von der Matrix
- Immer noch beliebige Verschachtelungen von Quantoren möglich
- Skolemisierung: Eliminieren der Existenzquantoren

Hintergrund

Mathematisch äquivalent:

1. für alle $x \in D$ gibt es ein $y \in D$, so dass $P(x, y)$ gilt
2. es existiert ein $f : D \to D$, so dass für alle $x \in D$ gilt $P(x, f(x))$

Problem!

Formalisierung obiger Aussagen ist nicht mehr FOL:

$$\forall x \exists y [P(x, y)] \equiv \exists f \forall x [P(x, f(x))]$$

Ansonsten könnte man damit Existenz- und Allquantoren trennen:

$$\begin{aligned} \forall x \exists y \forall u \exists v [P(u, v, x, y)] &\equiv \exists f \forall (x, u) \exists v [P(u, v, x, f(x))] \\ &\equiv \exists (f, g) \forall (x, u) [P(u, g(x, u), x, f(x))] \end{aligned}$$

Skolemisierung—Beispiel

- Existenzielle Quantifizierung über Funktionen ist bereits implizit im Begriff der Erfüllbarkeit enthalten
  - *Eine Formel ist erfüllbar, wenn eine Interpretation existiert, die sie erfüllt*

$\Rightarrow$ Entstehende Formel ist nicht mehr logisch äquivalent, jedoch noch erfüllbarkeitsäquivalent.

Beispiel (Skolemisierung)

$$\forall x \exists y \forall u \exists v [P(u, v, x, y)]$$

wird zu

$$\forall (x, u) [P(u, g(x, u), x, f(x))]$$

und wegen implizter Allquantifizierung der freien Variablen zu

$$P(u, g(x, u), x, f(x))$$

- f und g heißen Skolemfunktionen und dürfen in der bisherigen Formel nicht vorkommen

Skolemisierung—Korrektheit

Definition (Menge der Funktionen)

Die Menge $\mathtt{funcs}(\varphi)$ enthält alle Funktionssymbole der Formel φ.

Theorem (Korrektheit der Skolemisierung)

Ist φ eine Formel, mit

- $f \notin \mathtt{funcs}(\varphi)$ *und*
- $\mathtt{free}(\exists y[\varphi]) = \{x_1, \ldots, x_n\}$

dann gibt es zu einer beliebigen Interpretation $\mathcal{M}$ eine weitere Interpretation $\mathcal{M}'$, die sich nur in der Interpretation von f unterscheidet, so dass für alle Belegungen β_D gilt:

$$\mathtt{holds}(\mathcal{M}, \beta_D, \exists y[\varphi]) = \mathtt{holds}(\mathcal{M}', \beta_D, \varphi[y/f(x_1, \ldots, x_n)])$$

und ebenso

$$\mathtt{holds}(\mathcal{M}, \beta_D, \exists y[\varphi]) = \mathtt{holds}(\mathcal{M}', \beta_D, \exists y[\varphi]) \text{ , } \textbf{\textit{da}} \ f \notin \mathtt{funcs}(\varphi).$$

Skolemisierung—Vorgehen

- $\exists x$ Quantoren werden sukzessive eliminiert
- Verschiedene Strategien denkbar:
  - Von außen nach innen oder von innen nach außen eliminieren
  - Formel zuerst in Normalformen bringen?

Normalformen

- NNF ist ratsam, da man sonst möglicherweise überflüssige Skolemisierungen vollzieht: $\neg\exists x[x = 4] \equiv \forall x[x \neq 4]$
- PNF überflüssig, da sie möglicherweise mehr freie Variablen in den Scope eines existentiellen Quantors einführt

Skolemisierungs Reihenfolge

- Von außen nach innen eliminieren reduziert ebenfalls Anzahl freier Variablen:

$$\exists x \exists y[x \cdot y = 1] \rightsquigarrow \exists y[sk_0 \cdot y = 1] \rightsquigarrow sk_0 \cdot sk_1 = 1$$

versus

$$\exists x \exists y[x \cdot y = 1] \rightsquigarrow \exists x[x \cdot sk_0(x) = 1] \rightsquigarrow sk_1 \cdot sk_0(sk_1) = 1$$

Skolemisierung—Implementierung & Demo

Algorithmus: `skolemize(φ)`

Eingabe: Formel φ
Ausgabe: Matrix der Skolemisierung von φ

1. Berechne $\varphi' = \texttt{nnf}(\varphi)$
2. Skolemisiere in φ' Existenzquantoren von außen nach innen
3. Gib Matrix des Resultats zurück

Demo

```
phi1: ... = ∃(Y): (p(X,Y) → ∀(U) ∃(V): p(f(X,U),f(Y,V)))
phi2: ... = ∀(x): p(X) → ∃(X,Y): q(Y) ∨¬∃(Z): p(Z) ∧ q(Z)

phi1.skolemize
... = ¬p(X,sk0(X)) ∨ p(f(X,U0),f(sk0(X),sk1(U0,X)))

phi2.skolemize
... = ¬p(X) ∨ q(sk0) ∨¬p(Z0) ∨¬q(Z0)
```

Klauselnormalform

- Äquivalent zur CNF in der Aussagenlogik

Definition (Klauselnormalform)

Eine Formel ist in Klauselnormalform, wenn sie skolemisiert und in CNF ist.

Darstellung häufig als Klauselmenge und Literalmenge

Beispiel (Klauselnormalform)

$$(\forall x \exists y \forall u \exists v [P(u, v, x, y)] \vee R(z)) \wedge Q(z, w)$$

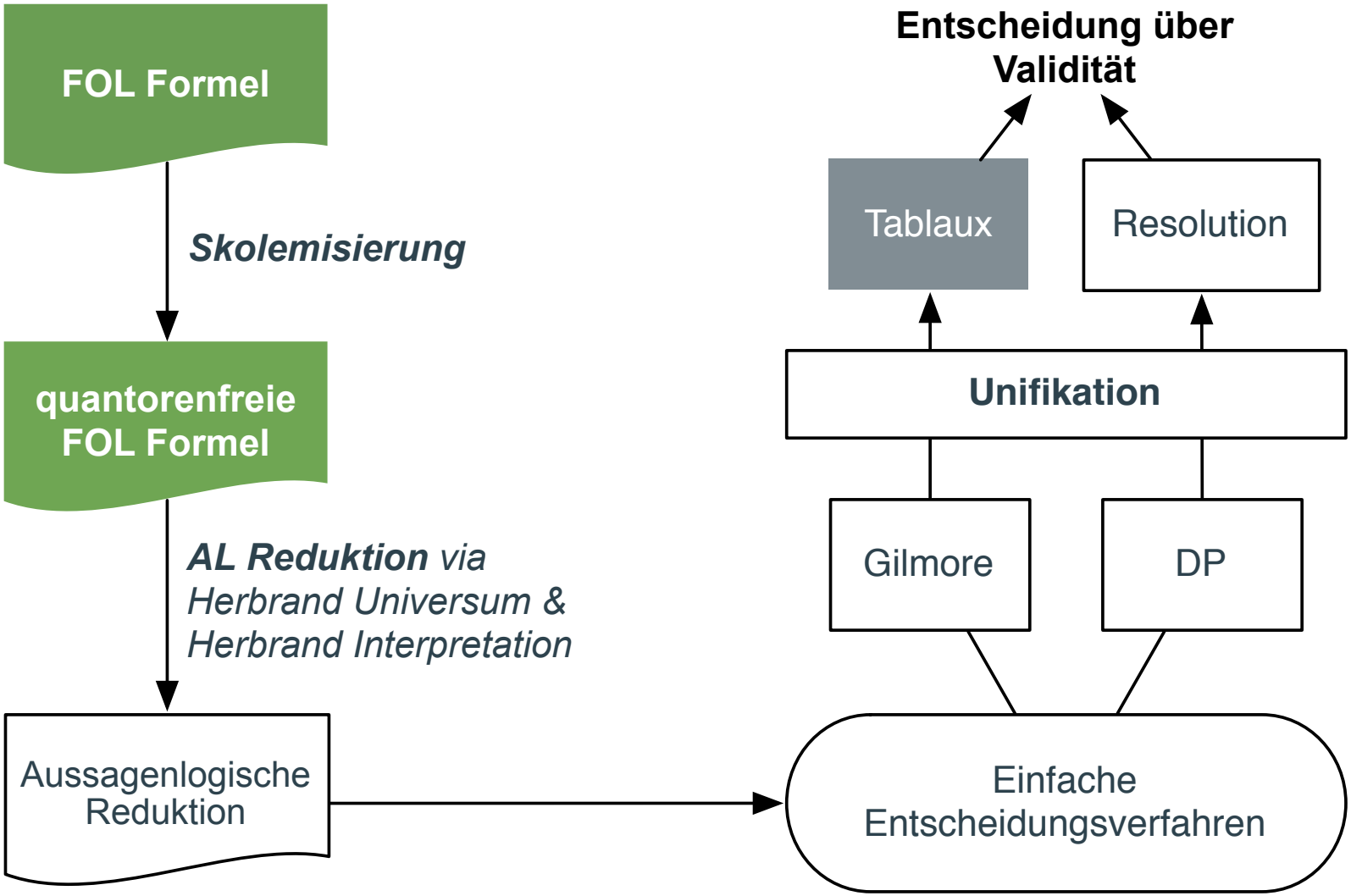
wird skolemisiert zu

$$(P(u, g(x, u), x, f(x)) \vee R(z)) \wedge Q(z, w)$$

und damit in Klauselnormalform zu

$$\{\{P(u, g(x, u), x, f(x)), R(z)\},\ \{Q(z, w)\}.$$

Wo stehen wir?



Kanonische Modelle

Quantorenfreie Formeln in FOL können auch aussagenlogisch betrachtet werden:

- Anstelle von aussagenlogischen Variablen haben wir Prädikatssymbole, die auf Terme angewandt werden
- Eine gegebene Formel hat immer eine endliche Anzahl an Variablen, Funktions- und Prädikatssymbolen
- Beweis der Kompaktheit kann angewandt werden

Idee!

FOL Formel:

$$P(a) \wedge R(x, y) \vee \neg R(x, f(x)) \wedge P(a)$$

AL Äquivalenz:

$$A \wedge B \vee \neg C \wedge A$$

? Interessante Frage!

Kann man die FOL-Validität zurückzuführen auf die aussagenlogische Validität?

Aussagenlogische Gültigkeit

- $\beta_{\mathcal{A}}$: Abbildung von atomaren Formeln auf Wahrheitswerte

Algorithmus: $\texttt{pholds}(\beta_{\mathcal{A}}, \varphi)$

Eingabe: $\beta_{\mathcal{A}}$ wie oben definiert, quantorenfreie FOL Formel φ
Ausgabe: `true`, wenn φ in aussagenlogischer Betrachtung gilt, sonst `false`

$\texttt{pholds}(\beta_{\mathcal{A}}, \varphi) = \varphi\ \texttt{match}$

| φ ist atomare Formel $\rightsquigarrow$ $\beta_{\mathcal{A}}(\varphi)$

| ... $\rightsquigarrow$ alle anderen Fälle wie bei $\texttt{eval}(\beta, \varphi)$

Beispiel (Aussagenlogische Gültigkeit)

- $\varphi = P(a) \wedge R(x, y) \vee \neg R(x, f(x)) \wedge P(a)$
- $\beta_{\mathcal{A}} = \{P(a) \mapsto \texttt{true}, R(x, y) \mapsto \texttt{false}, R(x, f(x)) \mapsto \texttt{false}\}$

$$\texttt{pholds}(\beta_{\mathcal{A}}, \varphi) = \texttt{true}$$

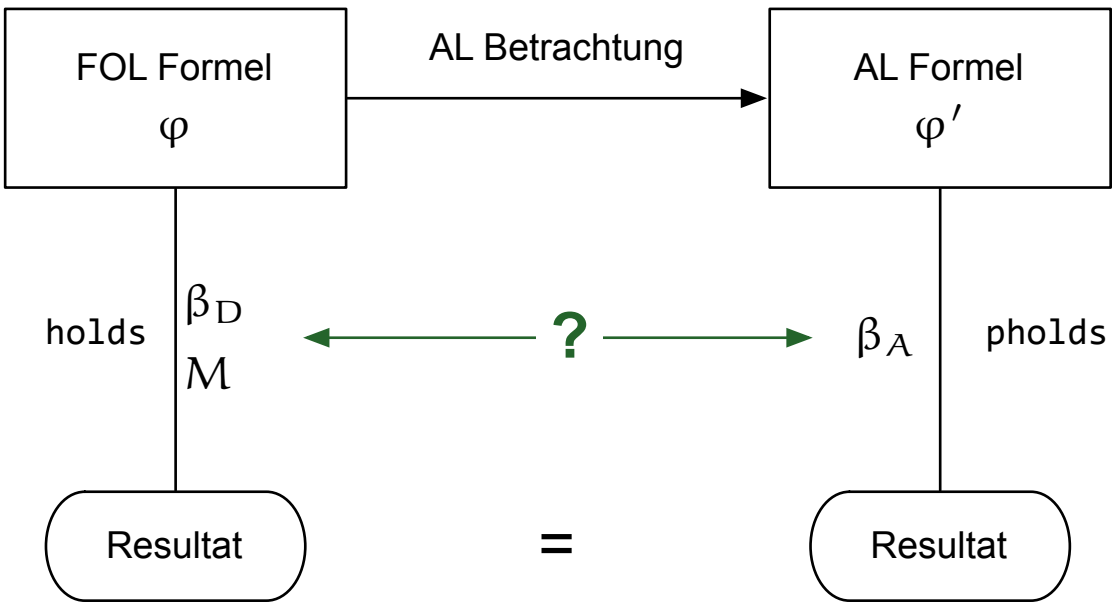
FOL Validität vs. AL Validität

Was wir zeigen wollen?

Eine quantorenfreie FOL Formel ist genau dann allgemeingültig, wenn die entsprechende aussagenlogische Betrachtung der Formel ebenfalls allgemeingültig ist.

Wie erreichen wir das?

Wir stellen eine Korrespondenz zwischen einer FOL Interpretation und Belegung und einer aussagenlogischen Belegung her.



Korrespondenz—1

$$\beta_D, \mathcal{M} \to \beta_A$$

- β_D und $\mathcal{M}$ definieren bereits eine entsprechende aussagenlogische Belegung

$$\beta_A(R(t_1, \ldots, t_n)) = \mathtt{holds}(\mathcal{M}, \beta_D, R(t_1, \ldots, t_n))$$

Theorem

Sei

- β_A *nach obigem Muster definiert*
- φ *eine quantorenfreie FOL Formel*

dann gilt für alle Interpretationen $\mathcal{M}$ *und alle Belegungen* β_D

$$\mathtt{pholds}(\beta_A, \varphi) = \mathtt{holds}(\mathcal{M}, \beta_D, \varphi)$$

Theorem

Wenn die AL Betrachtung φ' *einer quantorenfreien FOL Formel* φ *eine aussagenlogische Tautologie ist, so ist* φ *gültig.*

Korrespondenz—2

$$\beta_A \to \beta_D, \mathcal{M}$$

- Aus einer AL Belegung β_A wollen wir eine Interpretation $\mathcal{M}$ und Belegung β_D erzeugen, so dass gilt

$$\texttt{holds}(\mathcal{M}, \beta_D, \varphi) = \texttt{pholds}(\beta_A, \varphi)$$

- Vorgehen etwas technischer, aber der wichtige Punkt:

i Kanonische Interpretation

Wir können aus einer AL Belegung β_A eine Interpretation $\mathcal{M}_{\beta_A}$ (kanonische Interpretation) und eine Belegung β_D ableiten, so dass gilt:

$$\texttt{holds}(\mathcal{M}_{\beta_A}, \beta_D, \varphi) = \texttt{pholds}(\beta_A, \varphi)$$

Theorem

Eine quantorenfreie FOL Formel φ gültig, gdw. die AL Betrachtung φ' eine Tautologie ist.

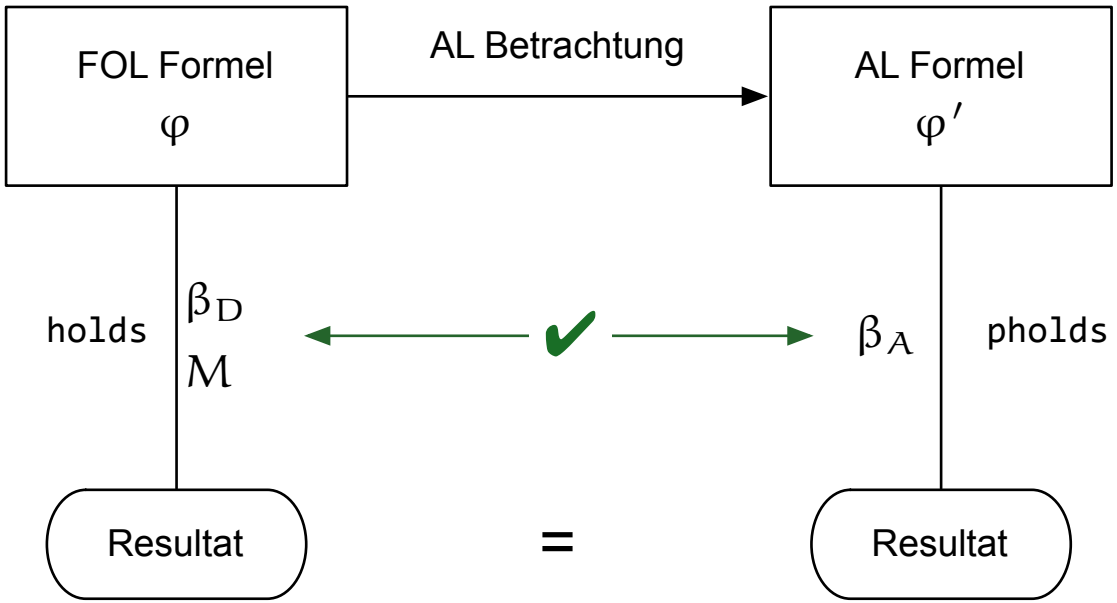
Was haben wir erreicht?

Was wir zeigen wollen?

Eine quantorenfreie FOL Formel ist genau dann allgemeingültig, wenn die entsprechende aussagenlogische Betrachtung der Formel ebenfalls allgemeingültig ist.

Wie erreichen wir das?

Wir stellen eine Korrespondenz zwischen einer FOL Interpretation und Belegung und einer aussagenlogischen Belegung her.



FOL Erfüllbarkeit vs. AL Erfüllbarkeit—1

- Die Korrespondenz zwischen FOL Validität und AL Validität ist interessant, aber:
  - Gilt nur für quantorenfreie Formeln
  - Skolemisierung (für quantorenfreie Formeln) ist nur erfüllbarkeits-erhaltend, nicht validitäts-erhaltend

i Variablenfreier Fall ist einfach!

Eine variablenfreie FOL Formel ist erfüllbar gdw. ihre AL Betrachtung erfüllbar ist.

Hinrichtung ist wieder einfach:

Theorem

Wenn eine quantorenfreie FOL Formel φ erfüllbar ist, so ist auch ihre AL Betrachtung φ' erfüllbar.

Wäre φ' nicht erfüllbar, so wäre $\neg\varphi'$ eine Tautologie. Damit wäre $\neg\varphi$ gültig und φ könnte nicht erfüllbar sein.

FOL Erfüllbarkeit vs. AL Erfüllbarkeit—2

Wir haben gezeigt: FOL Formel erfüllbar → AL Betrachtung erfüllbar

⚠ Umkehrrichtung ist nicht so einfach

Aussagenlogische Betrachtung von $P(x) \wedge \neg P(y)$ ist erfüllbar, nicht jedoch die original FOL Formel.

Wir gehen wieder den technischen Umweg über kanonische Interpretationen und daraus folgend kanonische Modelle.

Theorem

- φ *eine quantorenfreie FOL Formel*
- β_A *eine aussagenlogische Belegung von atomaren Formeln*
- $\mathcal{M}$ *eine kanonische Interpretation für* φ *mit* $P_{\mathcal{M}}(t_1, \ldots, t_n) = \beta_A(P(t_1, \ldots, t_n))$

dann gilt für eine beliebige Belegung β_D*:*

$$\mathtt{holds}(\mathcal{M}, \beta_D, \varphi) = \mathtt{pholds}(\beta_A, \mathtt{subst}(\beta_D, \varphi))$$

Und damit: AL Betrachtung erfüllbar → FOL Formel erfüllbar

Herbrand-Universum

Idee!

Finde ein kanonisches Modell mit kleinst-möglichem Universum.

Definition (Herbrand-Universum)

Das Herbrand-Universum zu einer FOL Sprache ist die Menge aller variablenfreien Terme dieser Sprache.

- Wenn die Sprache keine Konstanten hat, wird eine Konstante c hinzugefügt.
- Herbrand-Universum (HU) für eine Formel φ: HU für die Sprache aus φ

Beispiel (Herbrand-Universum)

- Sprache: $\{P^{(1)}, R^{(3)}, a^{(0)}, f^{(1)}\}$
- HU$(P(x) \wedge R(x,y,z) \vee \neg P(y)) = \{c\}$
- HU$(P(a) \wedge R(x,y,z) \vee \neg P(f(y))) = \{a, f(a), f(f(a)), \dots\}$

Herbrand-Interpretation

Definition (Herbrand-Interpretation)

Eine Herbrand-Interpretation ist eine kanonische Interpretation, deren Definitionsbereich das Herbrand-Universum ist.

D.h. eine Herbrand-Interpretation ordnet

- jedem Funktionssymbol f^n eine Funktion $f : HU^n \to HU$ zu

Beispiel (Herbrand-Interpretation)

- Sprache: $\{P^{(1)}, R^{(3)}, a^{(0)}, f^{(1)}\}$

$$\varphi = P(a) \wedge R(x, y, z) \vee \neg P(f(y))$$

- Herbrand-Universum

$$HU(\varphi) = \{a, f(a), f(f(a)), \dots\}$$

- Herbrand-Interpretation
  - $a_M = a$
  - $f_M(t) = f(t)$

Herbrand-Modell und Grund-Instanzen

Definition (Herbrand-Modell)

Ein Modell einer Menge Γ von Formeln, das eine Herbrand-Interpretation ist, heißt auch Herbrand-Modell.

Definition (Grund-Instanz)

Eine variablenfreie Formel $\varphi\sigma$ heißt Grund-Instanz von φ, wenn σ in das Herbrand-Universum abbildet.

Beispiel (Grund-Instanzen)

- Sprache: $\{P^{(1)}, R^{(3)}, a^{(0)}, f^{(1)}\}$

$$\varphi = P(a) \land R(x, y, z) \lor \neg P(f(y))$$

Mögliche Grund-Instanzen

- $P(a) \land R(a, a, a) \lor \neg P(f(a))$
- $P(a) \land R(f(a), a, a) \lor \neg P(f(a))$
- $P(a) \land R(a, f(a), a) \lor \neg P(f(f(a)))$
- $\ldots$

Der Satz von Herbrand

Theorem (Erfüllbarkeit in Herbrand-Modellen)

Eine Herbrand-Interpretation $\mathfrak{H}$ erfüllt eine quantorenfreie Formel φ gdw. sie die Menge aller Grund-Instanzen von φ erfüllt.

Dieses Ergebnis gilt nicht nur für die Erfüllbarkeit in einem speziellen Herbrand-Modell, sondern ganz allgemein für die Erfüllbarkeit:

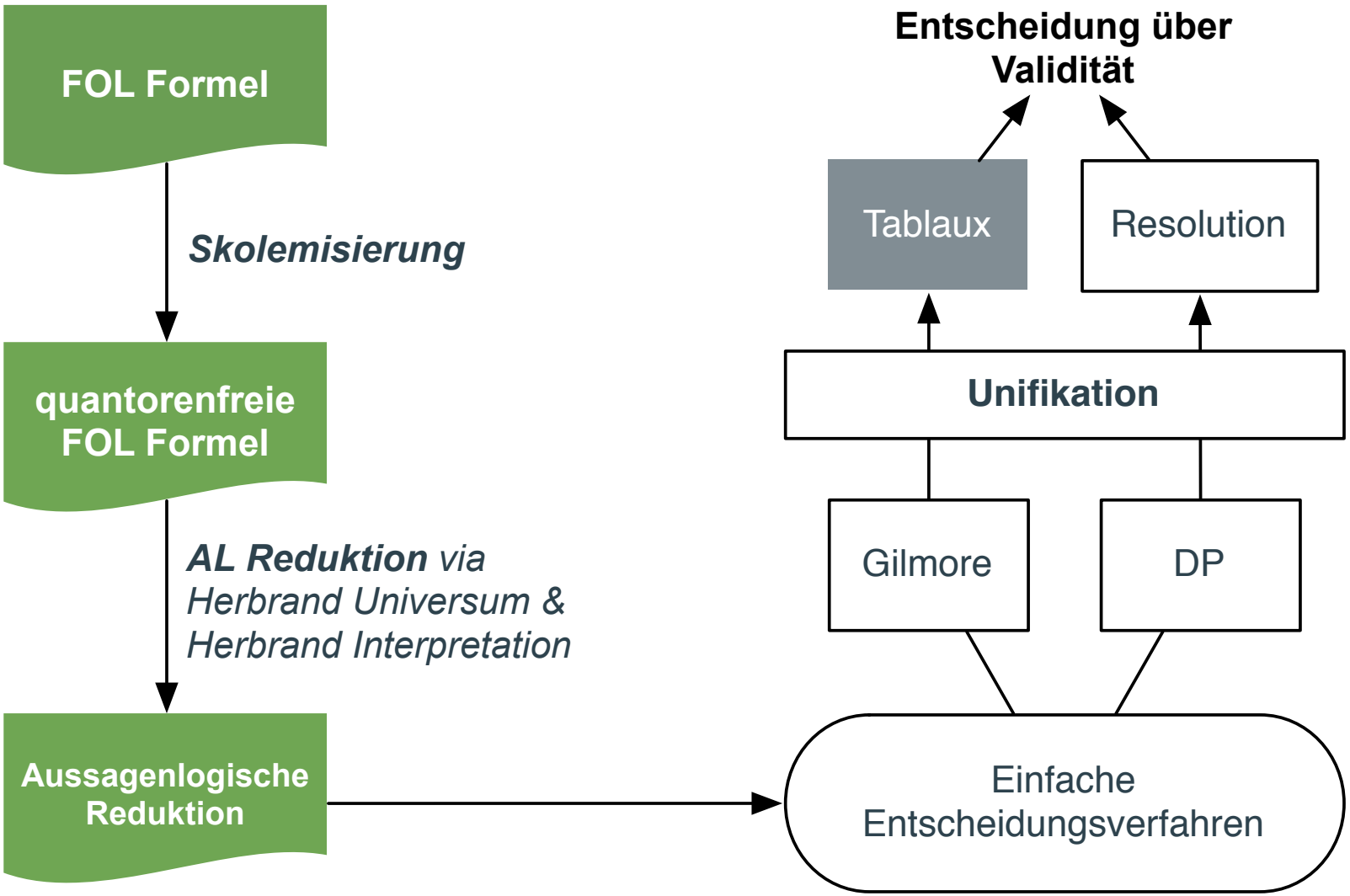
Theorem (Satz von Herbrand)

Eine quantorenfreie Formel φ ist erfüllbar in FOL, gdw. die Menge aller Grund-Instanzen aussagenlogisch erfüllbar ist.

Theorem

Eine quantorenfreie Formel hat ein Modell (d.h. ist erfüllbar) gdw. sie ein Herbrand-Modell besitzt.

Wo stehen wir?



Und wie implementiert man das?

Um die Erfüllbarkeit einer FOL Formel φ zu testen, können wir prinzipiell wie folgt vorgehen:

1. Bringe φ in Skolemform (und damit quantorenfrei bzw. rein universell quantifiziert)
2. Überprüfe, ob die Menge aller Grund-Instanzen aussagenlogisch erfüllbar ist

⚠ Problem!

Im Allgemeinen gibt es unendlich viele Grund-Instanzen!

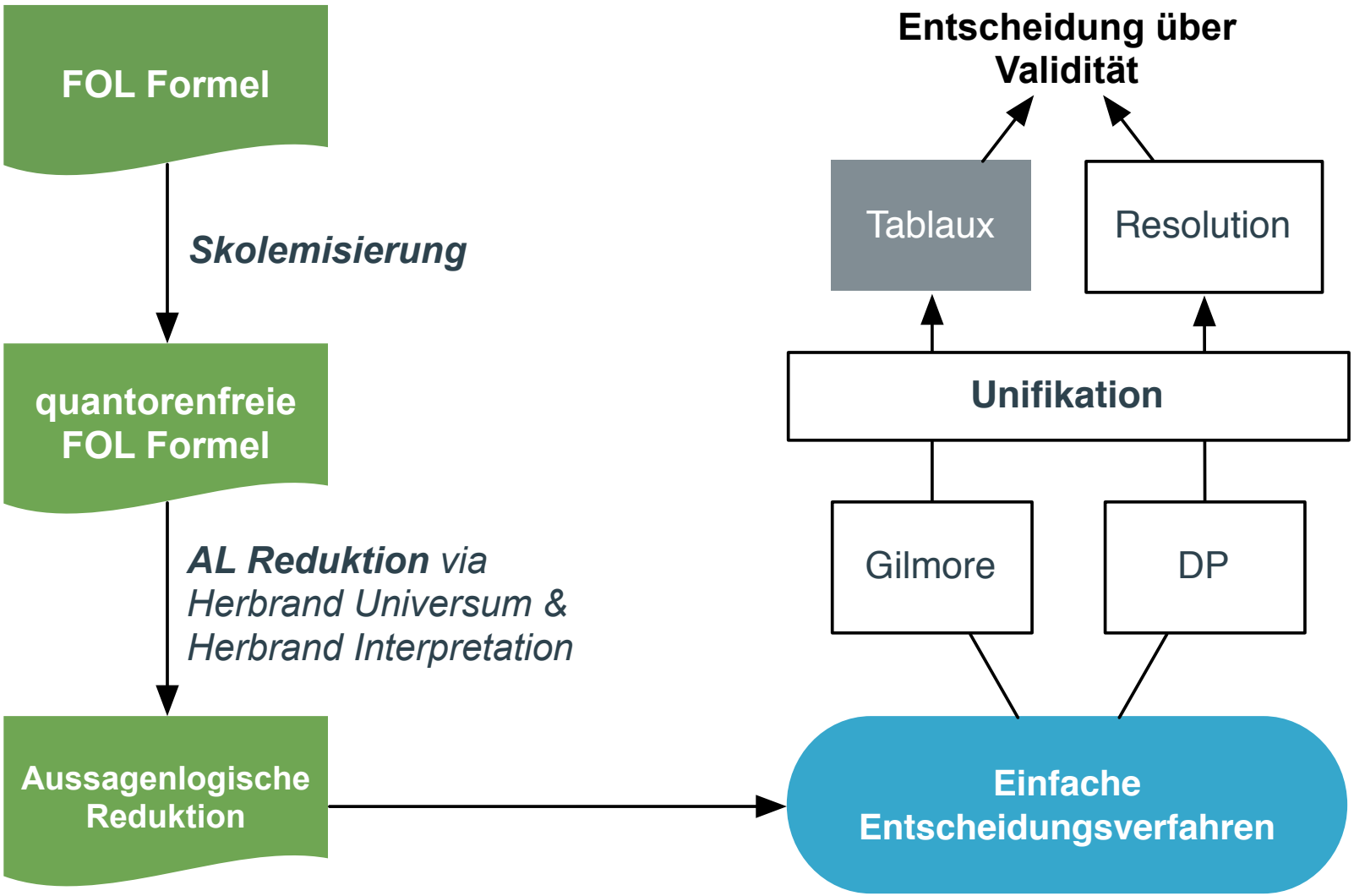
Idee!

Wende Kompaktheitssatz an: *Eine quantorenfreie FOL Formel ist erfüllbar, gdw. alle endlichen Mengen von Grund-Instanzen aussagenlogisch erfüllbar sind.*

Corollary (Unerfüllbarkeit einer FOL Formel)

Eine quantorenfreie FOL Formel φ ist unerfüllbar, gdw. eine beliebige endliche Menge von Grund-Instanzen aussagenlogisch unerfüllbar ist.

Wo stehen wir?



Ein erster Algorithmus von Gilmore (1960)

Grundidee!

1. Zähle immer größere Mengen von Grund-Instanzen auf
2. Teste die Konjunktion dieser Instanzen auf Unerfüllbarkeit

Verfahren von Gilmore (1960):

- In DNF konvertieren und komplementäre Literale suchen

Beispiel (Gilmores Verfahren)

- $\varphi = \exists x \forall y [P(x) \to P(y)]$
- Um Validität zu zeigen: $\varphi' = \neg \exists x \forall y [P(x) \to P(y)]$
- Skolemisierung: $\varphi'' = P(x) \wedge \neg P(sk_0(x))$
- 1. Grundinstanz: $i_1 = P(c) \wedge \neg P(sk_0(c))$ erfüllbar
- 2. Grundinstanz: $i_2 = P(sk_0(c)) \wedge \neg P(sk_0(sk_0(c)))$ erfüllbar, aber $i_1 \wedge i_2$ unerfüllbar

$\Rightarrow$ endliche Menge von Grundinstanzen ist unerfüllbar $\Rightarrow$ φ ist valide

Eine Verbesserung: Davis & Putnam (1960)

⚠ Problem bei Gilmore

- Man muss eine Konjuntion von Grund-Instanzen untersuchen
- DNF Konversion einer Konjunktion kann leicht explodieren

Bessere Idee!

Teste CNF auf Erfüllbarkeit, dann muss nur jede Grundinstanz selber in CNF konveritert werden, nicht jedoch das gesamte System.

Dies war der ursprüngliche Zweck der Davis Putnam Prozedur:

1. Zähle immer größere Mengen an Grundinstanzen auf
2. Überprüfe mit dem DP(LL) Verfahren, ob die entstehende Konjunktion erfüllbar ist oder nicht

i Vergleich

DP erzielt in der Praxis meist sehr viel bessere Ergebnisse als das Verfahren von Gilmore.

Wo stehen wir?