11 Verwendung von Theorien: Semantische Ansätze: Theorieresolution

Die Theorieresolution ist ein Schema, um Information über die Bedeutung von Prädikaten- und Funktionssymbolen unmittelbar im (Resolutions-) Kalkül auszunutzen, indem man anstelle von Axiomen für diese Symbole maßgeschneiderte Schlußregeln verwendet. Sie wurde von Mark Stickel am SRI allgemein formuliert. Viele Spezialfälle davon, u.a. Paramodulation, waren jedoch schon vorher unter anderen Namen bekannt.

Z.B. ist es dadurch möglich, direkt mit arithmetischen Bedingungen umzugehen, die man ohne diese Technik nur direkt (und umständlich) in Prädikatenlogik kodieren kann.

Zur Motivation der Vorgehensweise erinnern wir uns an die Begründung für die Korrektheit der einfachen Resolutionsregel:

$$\begin{array}{ll} \text{Klausel1:} & L, K_1, \ldots, K_n \\ \text{Klausel2:} & \underline{\neg L, M_1, \ldots, M_m \qquad\qquad} \\ \text{Resolvente:} & K_1, \ldots, K_n, M_1, \ldots, M_m \end{array}$$

Das entscheidende Argument dafür, daß die Resolvente aus den Elternklauseln folgt, war: wenn eine Interpretation das Literal L erfüllt, dann falsifiziert sie $\neg L$. Wesentlich ist also, daß keine Interpretation sowohl L als auch $\neg L$ erfüllen kann. Diese Eigenschaft haben zwei Literale dann, wenn sie die rein syntaktische Bedingung erfüllen, komplementär zu sein, also übereinstimmende Prädikatensymbole und Termlisten, aber verschiedene Vorzeichen zu besitzen.

In vielen Fällen kann man den syntaktischen Komplementaritätsbegriff verallgemeinern, indem man ausnutzt, daß nicht völlig beliebige, sondern nur bestimmte Klassen von Interpretationen in Frage kommen. Zum Beispiel könnte eine Formelmenge geeignete Axiome für das Prädikatensymbol $<$ enthalten, so daß nur solche Interpretationen Modelle sein können, die dem Symbol $<$ eine strikte Ordnungsrelation auf dem Universum zuordnen. Aufgrund der Eigenschaften solcher Relationen kann keine derartige Interpretation sowohl $a < b$ als auch $b < a$ erfüllen. In dem genannten Kontext sind diese beiden Literale zwar nicht syntaktisch, aber sozusagen semantisch komplementär, und auch folgender Ableitungsschritt ist korrekt:

$$\begin{array}{ll} \text{Klausel1:} & a < b, K \\ \text{Klausel2:} & \underline{b < a, M} \\ \text{Resolvente:} & K, M \end{array}$$

Als weitere Verallgemeinerung können wir sogar die Beschränkung auf zwei Elternklauseln aufgeben. Keine Interpretation der genannten Klasse kann sowohl $a < b$ als auch $b < c$ als auch $c < a$ erfüllen. Damit kann man analog zur Begründung für die einfache Resolution, nur mit mehr Fallunterscheidungen, auch folgenden Schritt als korrekt nachweisen:

$$\begin{array}{ll} \text{Klausel1:} & a < b, K \\ \text{Klausel2:} & b < c, M \\ \text{Klausel3:} & \underline{c < a, N} \\ \text{Resolvente:} & K, M \end{array}$$

Die Idee ist also, von dem Spezialfall zweier syntaktisch komplementärer Resolutionsliterale überzugehen zu einer beliebigen Menge von Literalen, so daß keine Interpretation einer gegebenen Klasse alle diese Resolutionsliterale erfüllen kann. Die "Klasse von Interpretationen" wird durch den Theoriebegriff präzisiert.

Einer erfüllbaren Menge $\mathcal{A}$ von Formeln läßt sich in PL_1 eindeutig die Klasse $\mathcal{M}$ ihrer Modelle zuordnen, also von Interpretationen, die die Formeln wahr machen. Dieser Klasse von Interpretationen entspricht wiederum eindeutig eine (unendliche) maximale Menge $\mathcal{T}$ von Formeln, die von allen Interpretationen in $\mathcal{M}$ erfüllt werden. Die Formelmenge $\mathcal{T}$ ist maximal in dem Sinn, daß jede weitere Formel die Modellklasse $\mathcal{M}$ einschränkt, also von wenigstens einem Modell für $\mathcal{A}$ falsifiziert wird. Per definitionem ist $\mathcal{T}$ gerade die Menge der Folgerungen aus $\mathcal{A}$. In dieser Sichtweise enthalten $\mathcal{M}$ und $\mathcal{T}$ dieselbe Information, und beide werden oft als Theorie bezeichnet. Da verschiedene Formelmengen dieselben Modelle haben können, stellt ein gegebenes $\mathcal{A}$ nur eine Alternative dar, die Theorie zu definieren. Man nennt $\mathcal{A}$ auch eine Präsentation oder Axiomatisierung der Theorie.

Definition 11.1. (Theorien)

Für eine Formelmenge $\mathcal{A}$ ist: Theorie($\mathcal{A}$) = Menge aller Modelle von $\mathcal{A}$ $\sim$ Menge aller Formeln, die in allen Modellen von $\mathcal{A}$ gelten.

Sei $\mathcal{T}$ eine gegebene Theorie:

$\mathcal{T}$-erfüllbar:: *Eine Formel F ist $\mathcal{T}$-erfüllbar, falls es ein Modell von $\mathcal{T}$ gibt, das F erfüllt.*

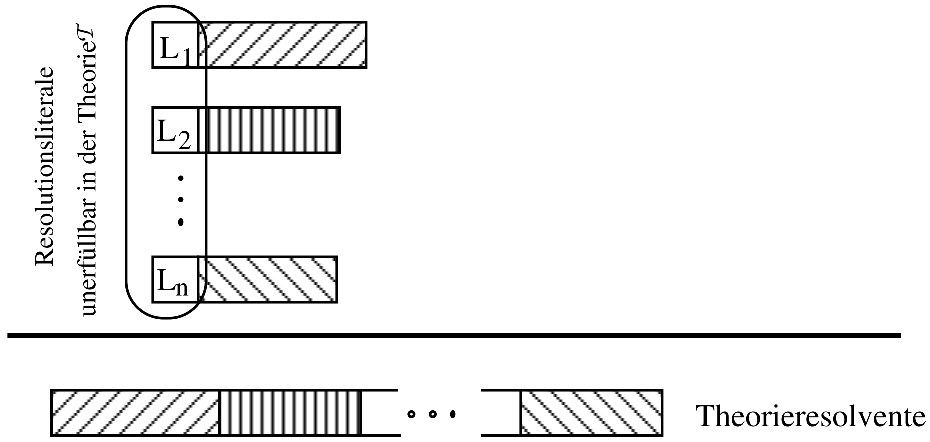
$\mathcal{T}$-unerfüllbar:: *Eine Formel F ist $\mathcal{T}$-erfüllbar, falls es kein Modell von $\mathcal{T}$ gibt, das F erfüllt.*

$\mathcal{T}$-allgemeingültig: *Eine Formel F ist $\mathcal{T}$-allgemeingültig: falls F in allen Modellen von $\mathcal{T}$ gilt.*

$\mathcal{T}$-Folgerung *$F \models_{\mathcal{T}} G$ Falls G in allen $\mathcal{T}$-Modellen von F gilt.*

Sei beispielsweise $\mathcal{A}$ die Menge $\{\forall x, y : P(x,y) \Rightarrow P(y,x)\}$, die nur aus dem Symmetrieaxiom für P besteht. Die Theorie $\mathcal{T}$ ergibt sich dann aus allen Interpretationen, die dem Prädikatensymbol P eine symmetrische Relation auf dem Universum zuordnen. Die Formel $P(a,b) \wedge \neg P(b,a)$ ist erfüllbar, aber nicht $\mathcal{T}$-erfüllbar. $P(b,a)$ ist eine $\mathcal{T}$-Folgerung von $P(a,b)$.

Definition 11.2. (Totale Theorieresolution ohne Unifikation) *Das aussagenlogische Schema für die totale Theorieresolution sieht nun folgendermaßen aus: Seien eine Theorie $\mathcal{T}$ und Klauseln $C_1, \ldots, C_n$ gegeben, und jede Klausel enthalte ein Literal L_i, so daß die Konjunktion all dieser Literale $\mathcal{T}$-unerfüllbar ist. Die Vereinigung der n Klauseln abzüglich dieser Resolutionsliterale L_i ergibt eine $\mathcal{T}$-Resolvente. Diese ist eine $\mathcal{T}$-Folgerung von $C_1 \wedge \ldots \wedge C_n$. Graphisch:*



Analog zur einfachen Resolution muß die Konjunktion der L_i im prädikatenlogischen Fall nicht unmittelbar $\mathcal{T}$-widersprüchlich sein. Man benötigt eine Substitution σ, einen $\mathcal{T}$-Unifikator, so daß die Formel $\sigma L_1 \wedge \ldots \wedge \sigma L_n$ $\mathcal{T}$-widersprüchlich ist. Die $\mathcal{T}$-Resolvente muß dann mit σ instantiiert werden. Allerdings ist ein allgemeinster $\mathcal{T}$-Unifikator für eine Menge von Ausdrücken nicht mehr bis auf Variablenumbenennung eindeutig bestimmt. Abhängig von $\mathcal{T}$ kann es einen, endlich viele oder unendlich viele voneinander unabhängige allgemeinste $\mathcal{T}$-Unifikatoren geben. Zwei Substitutionen werden dabei als unabhängig bezeichnet, wenn sich nicht eine durch Instantiierung von Variablen aus der anderen ergibt. In bösartigen Fällen existieren überhaupt keine allgemeinsten $\mathcal{T}$-Unifikatoren, sondern nur nicht-allgemeinste.

Eine Theorie mit endlich vielen $\mathcal{T}$-Unifikatoren ist die oben erwähnte Theorie der Symmetrie von P, die maximal zwei allgemeinste Unifikatoren erzeugt. Für

Klausel 1: $P(a,b), Q$

Klausel 2: $\neg P(x,y), R(x)$

besitzen die beiden ersten Literale die $\mathcal{T}$-Unifikatoren $\sigma_1 = \{x \mapsto a, y \mapsto b\}$ und $\sigma_2 = \{x \mapsto b, y \mapsto a\}$, so daß die zwei unabhängigen $\mathcal{T}$-Resolventen $\{Q, R(a)\}$ und $\{Q, R(b)\}$ abgeleitet werden können.

Das Konzept der Theorieresolution erlaubt, häufig vorkommende Standardinterpretationen von Symbolen wesentlich natürlicher und effizienter als mit der üblichen Axiomatisierung und normaler Resolution zu handhaben. Die Information über die spezielle Theorie steckt dabei in erster Linie im Unifikationsalgorithmus, der allerdings für jede Theorie neu entwickelt werden muß. Um die Widerlegungsvollständigkeit der Theorieresolutionsverfahren sicherzustellen, muß der verwendete Theorieunifikationsalgorithmus *alle* allgemeinsten Unifikatoren erzeugen bzw. im unendlichen Fall zumindest aufzählen können.

Der für die Realisierung der Theorieresolution erforderliche Unifikationsalgorithmus ist für manche Theorien zu aufwendig oder überhaupt nicht bekannt. Dies gilt insbesondere, wenn die Theorie eigentlich aus mehreren Untertheorien besteht, die aber nicht unabhängig voneinander sind. Als Beispiel betrachten wir die Theorie $\mathcal{T}_1$ der Interpretationen, die dem Prädikatensymbol $\leq$ (antisymmetrische) Ordnungsrelationen zuordnen, sowie die Theorie $\mathcal{T}_2$ der Interpretationen, die dem Prädikatensymbol $=$ die Gleichheitsrelation zuordnen. Die Kombination dieser beiden Theorien macht die Konjunktion der Literale $a \leq b, b \leq a, P(a), \neg P(b)$ unerfüllbar, so daß sie ohne weiteres als Resolutionsliterale für eine Theorieresolution in Frage kommen. Ein Algorithmus, der das erkennen kann, müßte jedoch genau für die Kombination dieser beiden Theorien konzipiert sein. Sobald eine dritte hinzukommt, ist er nicht mehr anwendbar. Deshalb versucht man, Algorithmen für einzelne Theorien zu entwickeln und einen allgemeineren Mechanismus verfügbar zu machen, der den Austausch zwischen beliebigen Theorien besorgt.

Betrachten wir die Kombination der genannten Theorien und folgende Klauseln:

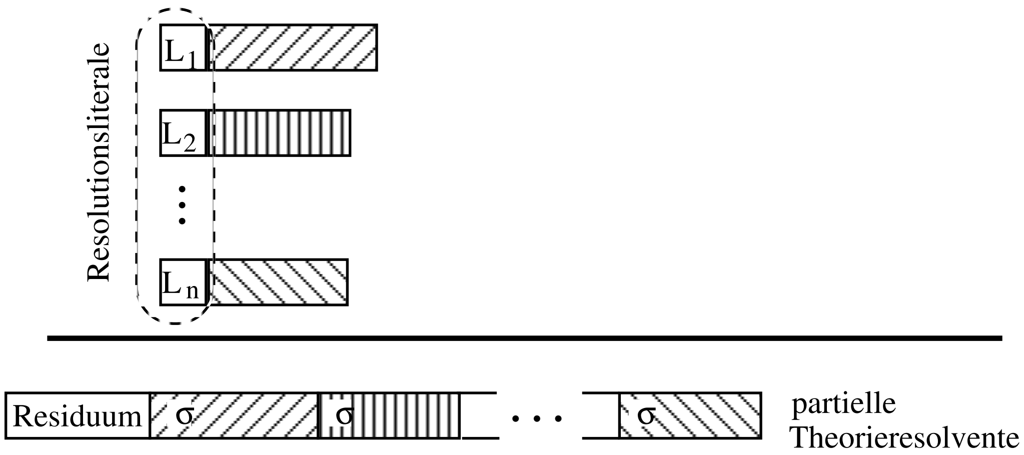
$$\begin{array}{ll} \text{Klausel1:} & a \leq b, \;\; K \\[1em] \text{Klausel2:} & b \leq a, \;\; L \\[1em] \text{Klausel3:} & P(a), \;\; M \\[1em] \text{Klausel4:} & \neg P(b), N \end{array}$$

Daraus müßte eigentlich die Resolvente $\{K, L, M, N\}$ ableitbar sein. Dies kann man durch einen verallgemeinerten $\mathcal{T}_1$-Schritt und einen anschließenden $\mathcal{T}_2$-Schritt erreichen. Wenn eine Interpretation in der Theorie $\mathcal{T}_1$ sowohl $a \leq b$ als auch $b \leq a$ erfüllt, dann muß sie wegen der Antisymmetrie auch das Literal $a = b$ erfüllen. Man sieht leicht, daß deshalb die Klausel $\{a = b, K, L\}$ eine $\mathcal{T}_1$-Folgerung aus Klausel 1 und Klausel 2 ist. Die Literale $a = b, P(a), \neg P(b)$ können dann von dem Algorithmus für $\mathcal{T}_2$ als Resolutionsliterale für eine Gleichheitstheorieresolution zwischen der neuen Klausel sowie Klausel 3 und Klausel 4 erkannt werden, womit die gewünschte Klausel $\{K, L, M, N\}$ entsteht.

Der erste Schritt geht über die Theorieresolution im bisherigen Sinn hinaus, da die Konjunktion der Resolutionsliterale allein noch nicht unerfüllbar in der Theorie ist, und da ein neues Literal abgeleitet und in die Resolvente aufgenommen wurde. Dieses sogenannte *Residuum* hat die Eigenschaft, daß es in der Theorie aus den Resolutionsliteralen folgt. Sein Prädikatensymbol braucht nicht einmal in den Elternklauseln vorzukommen. Ist ein Residuum beteiligt, spricht man von *partieller Theorieresolution*, andernfalls von *totaler Theorieresolution*.

Als Residuum kann man auch eine Disjunktion von mehreren Literalen zulassen. Das leere Residuum steht dann wie die leere Klausel für "falsch" und folgt daher nur dann aus den Resolutionsliteralen, wenn deren Konjunktion $\mathcal{T}$-unerfüllbar ist. Dieser Spezialfall entspricht der totalen Theorieresolution.

Definition 11.3. (Allgemeine (partielle) Theorieresolution) *Seien eine Theorie $\mathcal{T}$ und Klauseln $C_1, \ldots, C_n$ gegeben, und jede Klausel enthalte ein Literal L_i, so daß für eine Substitution σ gilt: $\sigma L_1 \wedge \ldots \wedge \sigma L_n \models_{\mathcal{T}}$ Residuum, dann ist $\sigma((C_1 \setminus L_1) \cup \ldots \cup (C_n \setminus L_n)) \cup$ Residuum eine allgemeine (partielle) Theorieresolvente.*
Schematisch:



Als Bedingung für die Widerlegungsvollständigkeit der partiellen Theorieresolution ergibt sich, daß der Unifikationsalgorithmus nicht nur alle allgemeinsten Unifikatoren, sondern auch alle "allgemeinsten Residuen" erzeugen können muß. Für das Suchverhalten des Verfahrens ist das sehr problematisch. Auch hier muß man heuristisch sich gute (kurze, kleine) Residuen aussuchen.

Paramodulation ist ein typisches Beispiel für partielle Theorieresolution. Die Theorie ist die Gleichheit, d.h. die Theorie, die sich aus dem Axiomensystem der Gleichheit (Definition 10.2) ergibt. Das paramodulierte Literal ist das Residuum.