

DIE ÜBERFÜHRUNG VON DER RELATIONALEN STRUKTURBESCHREIBUNG
IN DIE INTERFACE-STRUKTUR

Cornelia Zelinsky-Wibbelt
IAI/EUROTRA-D
Martin-Luther-Straße 14
D-6600 Saarbrücken

Der stratisfaktionelle Ansatz des EUROTRA-Projekts der Europäischen Gemeinschaft zur Maschinellen Übersetzung definiert zwischen Quell- und Zielsprache insgesamt 4 linguistisch motivierte Ebenen. Im vorliegenden Beitrag wird die Überführung von relationalen Strukturbeschreibungen (ERS) in die Repräsentation der Interface-Struktur (IS) behandelt. In diesem Schritt werden durch zusätzliche semantische Informationen lexikalische Mehrdeutigkeiten aufgelöst. Der Übersetzungszusammenhang zwischen ERS und IC muß durch T-Regeln beschrieben werden, deren Aufbau und Anwendung skizziert werden. An Beispielen wird u.a. gezeigt wie z.B. die Partikel oder Quantifikatoren der ERS beim Übergang auf die IS getilgt und durch semantische Merkmale ersetzt werden.

Anschließend an Schmidt 1987b wird im folgenden die Überführung von relationalen Strukturbeschreibungen auf die Ebene der Interface-Struktur beschrieben, die zusätzlich semantische Informationen einführt und so lexikalische Disambiguierungen vornehmen kann, die über die mit strukturellen Informationen erreichbaren hinausgehen; prinzipiell ist der Mechanismus derselbe wie beim Übergang von ECS nach ERS.

In den Beispielen dieses Beitrags wird die bisher bei EUROTRA als allgemeiner Standard geltende Numerierung der Argumente von Verben verwendet, die jedoch keinerlei semantische Konnotationen beinhaltet wie etwa Tiefenkasus, semantische Rollen, grammatische Funktionen im Sinne der LFG oder ähnliches. Eine Darstellung der Grundprinzipien der Interface-Struktur sowie die Vorschläge von EUROTRA-D zur feineren Differenzierung dieser Argumentbeschreibung finden sich in Steiner 1987c, die Vorstellungen zur systematischen Entwicklung der verwendeten semantischen Kategorien in Zelinsky-Wibbelt 1987b.

Bei der Erzeugung der semantischen Ebene geht es also um den Übergang von der dependentiellen Ebene, auf der die Relationen zwischen den syntaktischen Funktionen sowie ihre Werte in Form von syntaktischen Merkmalen spezifiziert werden, auf die Ebene, auf der die Relationen zwischen den semantischen Funktionen und ihren Werten - den semantischen Merkmalen - spezifiziert werden. Diese Übersetzung wird (wie beim Übergang ECS - ERS) durch einen modularen Apparat aus Übersetzungsregeln - genannt T-Regeln - erreicht. Jede T-Regel besteht aus einer linken und einer rechten Seite. Die funktionale Struktur auf der linken Seite der T-Regel bezieht sich auf eindeutige Repräsentationen, die auf der Quellebene erzeugt worden sind. Auf der rechten Seite werden die zu übersetzenden Funktionen spezifiziert. Im Falle von Tilgungen muß weiterhin auf der rechten Seite semantische Information über die getilgten Konstituenten spezifiziert werden. Die Generierung erfolgt durch rekursive Anwendung des Translators auf den Generator (vergl. Schmidt 1987b). Generator und Translator bestehen aus zwei Regeltypen: Strukturbauende Regeln eröffnen die Slots für die semantische Information in Form von semantischen Merkmalen, die den Prädikatsstellen und deren jeweiligen gesamten Potential von Argumentstellen zugeordnet sind. Diese Regeln werden Konstruktoren genannt.

Um einen hohen Grad an Generalisierung zu erreichen, wird die Struktur jeder syntaktischen Katego-

rie durch maximal einen Konstruktor gebaut. Dabei sind im Frame alle möglichen Argumente des Prädikats enthalten. Die vom Prädikat zu bindenden Argumentfunktionen sind alle fakultativ. Strukturfüllende Regeln - im Generator die atomaren B-Regeln, füllen die eröffneten Slots mit lexikalischer Information. Die lexikalische Information ist zum einen die Semantik einer jeden lexikalischen Einheit selbst, zum anderen werden die selektionalen Beschränkungen formuliert, die zu den Argumenten einer lexikalischen Einheit bestehen. Es wird deshalb in atomaren B-Regeln der maximale Frame deklariert, und die semantischen Rollen werden, wenn eine Restriktion in Form von semantischen Merkmalen ausgesprochen werden kann, dementsprechend markiert.

Die in den Konstruktoren durch variable Werte eröffneten Slots werden sukzessive mit den in den atomaren B-Regeln spezifizierten konstanten Werten aller Lesarten eines Wortlautes der betreffenden lexikalischen Kategorie gefüllt. Es erfolgt nur dann eine Bindung, wenn zwischen den semantischen Merkmalen die durch die Variablen geforderte Vertraglichkeit besteht. In diesem Fall werden beide Merkmalisten unifiziert.

Jede Unifikation bewirkt durch die Bindung der Funktionen die Erzeugung eines Ableitungsbaumes, dessen Auswertung eine Repräsentation ergibt. Sind alle Strukturebenen des Satzes ausgewertet, bewirkt die Unifikation auf der höchsten Ebene des Satzes die Erzeugung mindestens einer vollständigen Satzrepräsentation, das heißt, der Satz hat mindestens ein Übersetzungsergebnis.

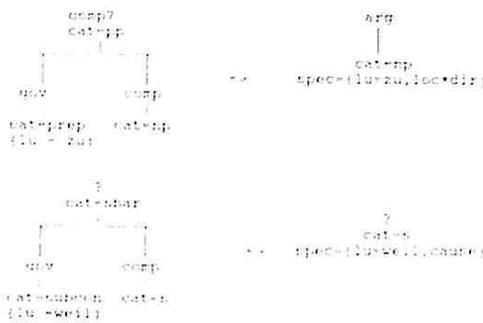
Generalisierungen über Relationen zwischen Attributen werden in A-Regeln ausgedrückt. Das betrifft das Verketten von semantischen Merkmalen in den Kopf aller Phrasen, sowie das Prüfen der Vertraglichkeit von Werten. Letzteres betrifft semantische Restriktionen, die zwischen einem Prädikat und seinen ungebundenen Elementen bestehen, wie zum Beispiel zwischen einem Substantiv und einer Adjektivphrase, die dieses modifizieren können:



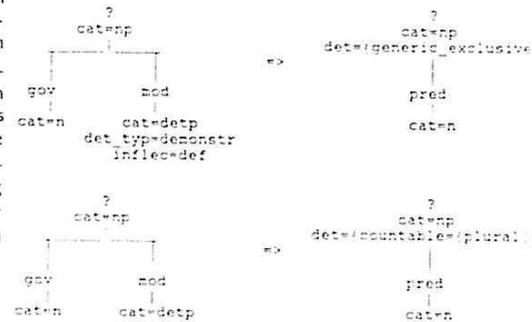
Regeln zur Formulierung der Semantik der möglichen Ergänzungen eines Prädikats

Die A-Regeln werden auf die gebauten Strukturen angewandt.

Im Übersetzungsprozess von der syntaktischen Dependenzebene in die semantische Ebene werden Generalisierungen über Attribute beim Übersetzen von Werten ausgedrückt. Partikel (valenzgebundene Präpositionen und Konjunktionen) werden im Übersetzungsprozess auf die semantische Ebene getilgt, wobei ihr Wortlaut als Wert des LU-Attributs in die Dekoration des Komplements - der NP bzw. des Satzes - kopiert wird und zusätzlich in semantische Merkmale übersetzt wird. Mit diesen beiden Informationen repräsentieren wir zur Zeit die Bedeutung der Präposition und der Konjunktion. Wir verdeutlichen diesen Übersetzungsprozeß an den folgenden Beispielen (in vereinfachter Form):



Auch die Determinantien und Quantifikatoren werden bei der Übersetzung in die Interface Struktur getilgt und in semantische Merkmale übersetzt:



Literatur

Die Literaturangaben für die Fachaufsätze zum Themenschwerpunkt Maschinelle Übersetzung finden sich zusammengefaßt auf Seite 23.