

Anaphernresolution in flach analysierten Texten für Recherche und Übersetzung

1 Einführung

Anaphernresolution ist wichtig für Retrieval und Textklassifikation, weil sie erlaubt, bei der Auszählung der Häufigkeit der Benennung eines Objekts, oder eines Objekttyps, eines Begriffs, die Ergebnisse durch die Anzahl der entsprechenden anaphorischen Verwendungen zu korrigieren. Bei der Maschinellen Übersetzung ist in der Regel durch diesen Bezug erst die korrekte Übersetzung von Pronomen möglich, aber auch bei anderen Kennzeichnungen leistet der Bezug eine wichtige Funktion bei der für die korrekte Übersetzung eventuell notwendigen Desambiguierung.

(1)

- a) *Paul kennt seine Frau seit langem; er ging sogar mit ihr zur Grundschule.*

Paul has known his wife for a long time; he even went to elementary school with her.

- b) *Weil er mit ihr in der Schule erwischt wurde, wurde Pauls nagelneue Wasserpistole beschlagnahmt.*

Because he was caught with it in school, Paul's brand new pistol was confiscated.

- c) *Der Regionalplan weist zwar eine Stelle für Landschaftsplanung aus, aber er sagt nichts über ihre Finanzierung.*

Although the regional plan shows a job for landscape management, it says nothing about its financing.

(2)

- a) *Maria hat sechs Glühbirnen gekauft. Sie will die Birnen in ihren neuen Leuchter einsetzen.*

Maria bought six light bulbs. She wants to put the bulbs in her new chandelier.

- b) *Maria hat ein Pfund blank polierte argentinische Birnen gekauft. Sie will die Birnen ihren Neffen schenken.*

*Maria has bought a pound of brightly polished Argentine pears. She wants to give the pears to her nephews.*¹

Der Kontrast zwischen (1.a) und (1.b) veranschaulicht die verschiedenen Übersetzungsmöglichkeiten des Personalpronomens *ihr* ins Englische, abhängig davon, ob es sich auf ein weibliches (menschliches) Wesen bezieht (Pauls Ehefrau in a)), oder nicht (auf die Pistole in b)). Wir sehen an diesem Beispiel auch, dass Pronomen sich nicht nur auf eine vorangehende NP beziehen können (anaphorisch), sondern auch auf eine nachfolgende (kataphorisch). Dass auch das Personalpronomen *er* verschiedene Übersetzungen ins Englische haben kann, abhängig davon, ob es sich auf ein männliches (menschliches) Wesen bezieht oder nicht, macht der Kontrast zwischen a) und c) deutlich, auch, dass die Possessivpronomen, hier *ihr* und *sein* dem nämlichen inhaltlichen Unterscheidungskriterium folgen. Bezogen auf andere Sprachpaare mögen die Verhältnisse anders liegen (Französisch unterscheidet wie das Deutsche nach grammatischem Geschlecht, zeigt beim Possessivpronomen aber im Gegensatz zu Deutsch und Englisch Kongruenz zum modifizierten Nomen.)

(2) zeigt wie die ambige NP *die Birnen* durch den Bezug auf *die Glühbirnen* einerseits und *das Pfund argentinischer Früchte* andererseits zu *the bulbs* bzw. *the pears* desambiguiert wird. Man beachte, dass diese Desambiguierung auf Grund des Satzkontexts allein nicht zu erhalten ist.

Seit langem ist bekannt, dass die gängigen morphosyntaktischen Kriterien ausreichen, um eine Mehrzahl der Personalpronomen in Texten korrekt auf ihre Antezedenten zu beziehen. Diese Kriterien sind morphologisch abzulesende Genus- und Numerus-Übereinstimmung und üblicherweise mehr oder weniger 'sophisticiertere' Verfeinerungen von syntaktischen Kommando-Konstellationen, die als Filterbedingungen fungieren können und Theorie-abhängig auf c-command (GB) oder o-command (HPSG) oder verwandten strukturell/funktionalen Kriterien beruhen. Schematisch: Ein Personalpronomen kann sich nicht auf eine NP beziehen, die das Pronomen lokal c-kommandiert, oder, mit HPSG-Hintergrund, lokal o-kommandiert, aber ein Reflexivpronomen kann das, und dieses nur in dieser Weise. Eine nicht-pronominale NP darf von seinem Antezedenten nicht nur nicht lokal sondern überhaupt nicht c- oder o-kommandiert werden, (vgl. die einschlägigen (Basis-) Studien dazu in Lasnik (1976), Chomsky (1981), Reinhart (1983), Pollard/Sag (1994), Lapin/McCord (1990) u.a.).

Diese Kriterien reichen aus, um beispielsweise die Pronomen in (1.a) nicht falsch beziehen zu können, außer *ihre* in c), das sich danach sowohl auf *eine Stelle* als auch auf *Landschaftspflege* beziehen kann, zumindest wenn Letzteres als referentiell verstanden wird.

¹ Die Beispiele stammen aus Testsuites zu *Personal Translator* und gehen zum Teil auf die Arbeiten von Ulrike Schwall und Herb Leass zurück.

In der Regel erzielt man gute Erfolge damit, wenn man die Auswahl aus der nach morphosyntaktischer Filterung verbleibenden Menge von Antezedent-Kandidaten auf Präferenzheuristiken stützt, wie sie sich aus den strukturellen Gegebenheiten ablesen lassen, die Aufmerksamkeitsregularien abbilden: Man weiß, dass neu eingeführte Diskursreferenten eher aufgegriffen werden als ältere bisher nicht wieder aufgegriffene, dass der Diskursreferent (DRF) aus dem Subjekt in der Regel eher als thematische Basis dient, als ein DRF aus den mehr 'obliquen' Komplementen oder Adjunkten, dass allgemein Diskursreferenten aus dem Hauptsatz (thematisch) stärker wahrgenommen werden als andere und insofern jeweils eher als Antezedent gewählt werden. Gleichzeitig spielen strukturelle (und inhaltliche) Parallelitäten zwischen Anapher und möglichem Antezedent eine große Rolle.

Das richtige Mischungsverhältnis solcher Kriterien zu bestimmen ist nicht leicht, zumal verschiedene Texte (Textsorten) und Domänen und Sprachen unterschiedliche Anforderungen an Trennschärfekapazitäten stellen und die Evaluierung von Verfahren keine leichte Aufgabe ist. Grosso modo kann man sagen, dass Verfahren, die die beschriebenen Kriterien angemessen verwenden, eine Güte von nahe 80 % der Fälle erzielen, im General Purpose-Bereich wohl gemerkt, nicht angepasst an eine spezielle Domäne oder Textsorte und bezogen nur auf die Auflösung von Pronomen. Dies sind jedenfalls die Werte die wir aus langjähriger Erfahrung mit dem Verfahren von Lappin/McCord (1990) bzw. Lappin/Leass (1994) (unter Einbezug von Varianten) gewonnen haben und sich im Wesentlichen mit Erfahrungswerten aus der Literatur zu ähnlichen Verfahren decken (vgl. zu Verfahrenstypen, Präferenzkriterien, Evaluation die Übersicht in Mitkov (1999)).

Diese Verlässlichkeit ist nicht ausreichend, wenn man hohe Qualität anstrebt, zumal im Bereich der Maschinellen Übersetzung, dann wenn man sich nicht auf spezielle Domänen einschränken will und vor allem auch dann, wenn man die Ergebnisse der Resolution maximal zur Textdesambiguierung nutzen will, weil sich dann einzelne Fehler wellenförmig ausbreiten können, vgl. Beispiel (3).

- (3) *Den Bewerber sieht Peter schließlich wieder beim Laserdrucker.*
 a) *Er soll ihn interviewen und, wenn er ihm gefällt, ihn einstellen*
 b) *Er soll ihn reparieren und wieder einstellen.*

Finally, Peter sees the applicant again at the laser printer.
 a) *He shall interview him and, if he likes him, hire him.*
 b) *He shall repair it and adjust it again.*

Dass das Subjekt der alternativen Folgesätze a) und b) sich auf *Peter* bezieht, erhält man in der Regel aus den Parallelitätserwartungen auch ohne weitere semantische Evaluation (die natürlich aber auch diese Entscheidung gegenüber der Alternative

Bewerber noch stärker legitimiert). Genauso, was die jeweiligen Wiederaufnahmen des Objekts in b) und im Hauptsatz von a) betrifft. Dass Parallelität auch bezüglich des wenn-Satzes in a) gewahrt ist, erhält man aber nur durch den Bezug auf die den syntaktischen Funktionen zugeordneten semantischen Rollen, bzw. aus dem Wissen über die passivische thematische Strukturierung des (syntaktisch) aktivisch verwendeten Verbs *gefallen*. Und ob sich die Kette der Pronomen *ihn, er, ihn* in a) auf den Bewerber oder den Drucker bezieht, ist nur sinnvoll zu entscheiden, wenn man weiß, dass Interviews mit Menschen, nicht Instrumenten geführt werden. Ähnlich in b), wo das Wissen darüber, dass Instrumente, nicht Menschen repariert werden, den Ausschlag für die umgekehrte Entscheidung gibt. Von dieser Entscheidung ist aber nicht nur abhängig ein, sondern gleich mehrere Vorkommen von *ihm/ihn* durch *him* bzw. *it* richtig oder eben falsch zu übersetzen, und, abhängig wiederum davon, auf Grund der semantischen Selektionserwartungen des ambigen Verbs *einstellen*, dieses im Sinne von *hire* oder aber im Sinne von *adjust* zu interpretieren.

Es liegt also auf der Hand, semantisches Wissen zu benutzen und einen semantischen Filter zu definieren, zumal wenn sich die zu definierende Resolutionskomponente in ein Gesamtsystem integriert, das ohnehin über solches Wissen schon verfügt oder prinzipiell verfügen muss. Semantische Kriterien sind mit Erfolg für die Resolution benutzt worden, vor allem durch das Ausnutzen von semantischen Selektionsbeschränkungen (vgl. Guenther/Lehmann (1983), Carbonell/Brown (1988), Dagan/Itai (1990), Connolly et al. (1994), Nakaiwa/Ikehara (1995), Mitkov (1998b)) zu Anwendungen in verschiedenen Verfahrenstypen).

Seit langem bekannt ist, dass nicht wenige Fälle von Referenz nur vernünftig zu erfassen sind, wenn die inhaltliche Strukturierung des Textes im Sinne von Repräsentationstheorien wie der *file change semantics* (Heim (1983)) oder der *Diskursrepräsentationstheorie* (DRT, Kamp (1981)) vorliegt. Denn dann sind Modaleinbettungen, Quantifikationsstrukturen und dergleichen hierarchisch beschrieben und erlauben vorherzusagen, ob ein bestimmter Antezedent-Kandidat für einen referentiellen Ausdruck überhaupt erreichbar, *zugänglich* ist, oder, ob es statthaft ist, einen als Antezedenten benötigten Plural-Referenten zu kreieren etc.²

Allerdings ist Voraussetzung für den Erhalt solch reichhaltiger semantischer Constraints, dass die Repräsentationen der Sätze zu spezifischen Lesarten mit festgelegten Skopus- und Applikationsverhältnissen ausdifferenziert sind. Genau solche Ausdifferenzierung will man gemeinhin beim Design linguistischer Software vermeiden, aus Performanzgründen und auch weil eine fundierte Entscheidung über die Les-

² Zur Definition von und zu Erfahrungen mit Verfahren, die auf Diskursrepräsentationsstrukturierung und Zugänglichkeitsrelation fußen vgl. Guenther/Lehmann (1983), Bosch/Geurts (1990), Sandt (1992), Kamp/Reyle (1993)].

art im Allgemeinen sehr viel linguistisches Wissen, noch mehr Weltwissen und komplexe Entscheidungsmechanismen voraussetzt.

Das im Folgenden vorgeschlagene Resolutionsverfahren versucht in dieser Hinsicht eine Optimierung. Es nutzt semantische und pragmatische Information zur Verbesserung der Filterung. Es operiert aber nicht über einer ausdifferenzierten semantischen Repräsentation, sondern über einer Menge semantischer Teilanalysen, die jedoch bei Bedarf durch das Resolutionsverfahren entsprechend der aktuellen syntaktischen und semantischen Constraints sukzessive spezifiziert werden können. Damit ist insbesondere a priori keine vollständige Zugänglichkeitsrelation im Sinne der DRT spezifiziert und semantische Beschränkungen, wie sie etwa aus lexikalischen Selektionsbeschränkungen folgen, sind nicht global, sondern nur lokal propagiert. Das Performanzproblem der Konstruktion einer spezifischen Lesart wird dadurch ökonomisch vermieden, zu dem Preis, dass semantische Information nur relativ lokal zur Verfügung steht. Das Verfahren kann aber, auf Grund einschlägiger Phänomen-Pattern oder wenn die Präferenzgewichte der Antezedentkandidaten zu dicht beieinander liegen, dynamisch zusätzliche Auswertung anstoßen, um dadurch stärkere Filterung bzw. signifikantere Gewichtung zu erzielen. Umgekehrt kann die Wahl eines Antezedenten bestimmte Festlegungen der Zugänglichkeitsrelation oder der semantischen Beschreibung des Antezedenten oder der Anapher bedingen und damit eine partielle Spezifizierung der Lesart.

Der Algorithmus erweitert das in Lappin/McCord (1990) und Lappin/Leass (1994) bzw. Leass/Schwall (1991) vorgeschlagene Resolutionsverfahren für syntaktische Analysen der Slot Grammar (McCord 1991), das sich, wie skizziert, auf morphosyntaktische Filter (mit Slot Grammar spezifischer Bindungstheorie) und dynamischer Gewichtung der Antezedent-Kandidaten stützt. Der Algorithmus inspiziert die aus der syntaktischen Analyse abstrahierte semantische Repräsentation, deren Expressivität der einer flachen unterspezifizierten Diskursrepräsentationsstruktur (FUDRS) entspricht (dh. die im Wesentlichen aus DRSen zu Teil-Phrasen mit flacher lexikalischer Semantik und Constraints über Applikation und DRF-Identifikation besteht, vgl. Kamp/Reyle (1993) zu DRT, Reyle (1993) zu UDRT, Eberle (1997) zu FUDRT). Das Verfahren ist zu einem großen Teil implementiert und in der Übersetzungssoftware *Personal Translator* enthalten (vgl. <http://www.linguatec-es.de>).

Wir beschränken uns im Folgenden auf die Wirkungsweise bei der Pronomenresolution, weisen aber darauf hin, dass für die Auflösung von Kennzeichnungen semantische Information eine noch größere Rolle spielt.

2 Lokale semantische Constraints

2.1 Selektionsbeschränkungen

Die Lexika des Systems differenzieren ihre Lemmata nach Valenz und Mehrdeutigkeit in verschiedene (Wort-)Bedeutungen die semantisch bezüglich einer semantischen Hierarchie von Nomen- Verb-, Adjektiv-, Adverb- und Präpositionsbedeutungen klassifiziert sind. Daneben ist relationale Information vorhanden, die spezifiziert von welchem Typ ein bestimmtes Komplement eines Verbs oder Nomens sein muss, aber auch welche Ereignisse oder Objekte typischerweise von einem bestimmten Adverb- oder Adjektivtyp modifiziert werden und, vor allem, von welcher Art typischerweise der Füller der POSS-Rolle einer Nomen-Bedeutung ist. Diese Information gibt fundierte Auswahlkriterien für Fälle wie die folgenden an die Hand:

- (4) *Die Hunde entdeckten die Kekse.*
 - a) *Sie fraßen sie.*
 - b) *Sie schmeckten lecker.*
- (5)
 - a) *Die Frau, die Probleme und ihre Idee.*
 - b) *Die Frau, die Probleme und ihre Lösung.*

Wenn wir wissen, dass das Subjekt von *fressen* vom Typ ANIMAL ist und das Objekt vom Typ FOOD, und *Hund* und *Keks* entsprechende Typinformation tragen, ist die Resolutionsaufgabe aus (4.a) gelöst. Ähnlich die durch (5.a) gestellte Aufgabe, wenn wir wissen, dass der 'Besitzer' von Ideen gemeinhin etwas Vernunftbegabtes ist (SENT), dann kann sich *ihre* nur auf *die Frau*, nicht aber auf *die Probleme* beziehen. (5.b) zeigt Grenzen auf und macht deutlich, dass die Rollen, zumindest solche, wie das unspezifische POSS nicht eindeutig zu sein brauchen. POSS zeigt das bekannte Ambiguitätsproblem bei relationalen Nomen, Agens des mit dem Nomen assoziierten Ereignisses zu sein oder Komplement des Nomens. Damit trifft die semantische Evaluierung hier (zurecht) keine Entscheidung und Darstellungsgesichtspunkte wie relative Nähe etc. müssen die Entscheidung bringen. (4.b) thematisiert ein anderes Problem: Hier ist die Entscheidung zugunsten *die Kekse* manifest, wenn wir wissen, dass das Subjekt semantisch nicht Agens, nur Thema ist. Das kann aber für die Prädikats-Argument-Struktur aus dem Beitrag des (entsprechend) getypten Adjektivs geschlossen werden (*lecker* modifiziert FOOD und ist damit nicht adverbial, sondern prädikativisch das Subjekt charakterisierend zu interpretieren).

2.2 Selektionsrestriktionen sind Präsuppositionen

(6) macht explizit, was (4.b) schon assoziiert hatte: den bekannten Sachverhalt, dass Rollenerwartungen Präsuppositionen sind. In unserem Zusammenhang bedeutet dies vor allem, dass die entsprechenden Constraints 'Negations-stabil' sind.

- (6) *Die Hunde entdeckten die Kekse.*
 a) *Sie waren nicht knusprig.*
 b) *Sie waren hungrig.*

Obwohl die Prädikation in (6.a) negiert ist, ist sie in ihrer Erwartung, was die Art des Subjekts betrifft genauso eindeutig, wie (6.b), so dass die korrekte Auflösung klar entschieden werden kann.

2.3 Ableitungen

Die präsuppositionale Erwartung kann ausgenutzt werden, um bei einem ansonsten den Kriterien genügenden Antezedent-Kandidaten bei seiner Wahl auf Eigenschaften zu schließen, die direkt nicht gegeben sind und diese für das (weitere) referentielle Textverstehen auszunützen. Das System nutzt solche Ableitungen für die Bestimmung von nichtspezifiziertem (grammatischem und auch natürlichem) Genus:

- (7) *Les journalistes disent la vérité et elles mentent.*
Die Journalistinnen sagen die Wahrheit und sie lügen.

Französisch *les journalistes* kann sowohl als *die Journalisten*, als auch als *die Journalistinnen* übersetzt werden. Das Genus der NP ist zunächst nicht spezifiziert. Damit ist diese als Antezedent des Feminin-Pronomens erlaubt. Wird sie als Antezedent gewählt, dann überträgt *elles* seine Genus-Ausprägung auf sie und erlaubt die korrekte feminine Übersetzung.

Die Informationsquelle kann auch außerhalb des Pronomens liegen und damit die Desambiguierung der Anapher gleich mit auslösen:

- (8) a) *Aldi kommt mit superbilligen neuen Produkten. Er ist Marktführer.*
Aldi comes with super cheap new products. It is market leader.
 b) *Aldu verführt seine Frau. Er ist ein Schlingel.*
Aldu seduces his wife. He is a rascal.

Auch auf der Basis großer Lexika sind Namen häufig nicht bekannt (auch weil laufend neue, bspw. für Firmen, kreiert werden). Andererseits ist es für Textklassifikation und

Übersetzung mitunter sehr wichtig zu wissen, auf welche Art von Objekt sich der Name bezieht. (8) macht das am Beispiel der Übersetzung ins Englische deutlich, bei der für die pronominale Wiederaufnahme gewusst werden muss, ob sich der Name auf einen Menschen bezieht oder nicht. Annahmen über die Rollenbesetzungen einschlägiger Verben, in diesem Fall durchaus auch Domänen-spezifisch, können es dann erlauben, wie in (8.b) darauf zu schließen, dass ein Name sich auf einen Menschen bezieht und ein darauf Bezug nehmendes *er* (bzw. jedes weitere solche Pronomen), dann mit *he* übersetzt werden muss, bzw., dass der Name sich, wie in (8.a), auf eine Firma bezieht, so dass das referierende Pronomen mit *it* zu übersetzen ist. Der Exegese des ersten Satzes kann man im Fall von (8.b) sicherlich widersprechen, auch wenn man von einer sehr eingeschränkten Domäne ausgeht, und man kann behaupten, dass *mit superbilligen Produkten kommen* keinen Subjekts-Typ FIRMA oder INSTITUTION erzwingt (sondern auch einen Menschen als Agens erlaubt). Aber der zweite Satz deutet ebenfalls, wenn nicht noch stärker, auf diese Interpretation hin und beide Informationen zusammen favorisieren in jedem Fall diesen Bezug.

Damit können wir, diesen Teil abschließend, den Präsuppositionsconstraint formal folgendermaßen beschreiben:

$$\begin{array}{l}
 \forall \text{QUAL1, QUAL2, CONTEXT, ANTE, PRON :} \\
 \left. \begin{array}{l}
 \text{CONTEXT} \vdash_{\text{presup}} \text{ANTE} \in \text{QUAL1} \ \& \\
 \text{CONTEXT} \vdash_{\text{presup}} \text{PRON} \in \text{QUAL2} \ \& \\
 \neg (\text{SEM_HIERACHY} \ \& \ \text{CONTEXT} \vdash \text{PRON} \notin \text{QUAL1}) \ \& \\
 \neg (\text{SEM_HIERACHY} \ \& \ \text{CONTEXT} \vdash \text{ANTE} \notin \text{QUAL2}) \ \&
 \end{array} \right\} \\
 \Rightarrow \\
 \left. \begin{array}{l}
 \text{possible}_{\text{presup}} (\text{antecedes}(\text{ANTE}, \text{PRON})) \ \& \\
 (\text{antecedes}(\text{ANTE}, \text{PRON}) \rightarrow \text{ANTE}, \text{PRON} \in \text{QUAL1}, \text{QUAL2}) \ \& \\
 (\text{SEM_HIERACHY} \vdash (\exists \text{QUAL3} [\text{QUAL1} \cup \text{QUAL2} \text{ relevant}(\text{QUAL3})]) \\
 \rightarrow \text{bonus}_{\text{presup}} (\text{antecedes}(\text{ANTE}, \text{PRON}))
 \end{array} \right\}
 \end{array}$$

Will heißen, falls für das Pronomen und einen möglichen Antezedenten Eigenschaften QUAL1 bzw. QUAL2 präsupponiert werden können, die nicht im Widerspruch zum Typ-Wissen über Antezedent und Pronomen stehen, dann können unter Präsuppositionsgesichtspunkt diese Diskursreferenten aufeinander bezogen werden und wenn dies geschieht, dann kann für beide DRFs von den Eigenschaften ausgegangen werden, die für wenigstens einen präsupponiert wurden. Wenn sie gar eine solche Eigenschaft von vornherein gemeinsam haben, die relevant ist, dann erhält der Bezug unter Präsuppositionsgesichtspunkten sogar einen Bonus, wobei 'relevante Eigenschaft' genügend spezifisch durch das System prädefiniert sein muss.

Boni werden Auflösungsmöglichkeiten auch dann zuerkannt, wenn aus den semantischen Typen der beteiligten Verben und NPs auf parallele Besetzung von semanti-

schen Rollen durch möglichen Antezedent und Pronomen geschlossen werden kann und entsprechend ein Malus, wenn auf das Gegenteil geschlossen werden kann.

3 Partielle Zugänglichkeitsrelation

Durch die Kombination lexikalischer und einer Reihe struktureller Ambiguitäten, u.a. verursacht durch Skopus tragende Elemente, wie Negations-, Modal-, Quantifikationsoperatoren haben Sätze in der Regel eine sehr hohe Anzahl von Lesarten aus denen NLP-Systeme, anders als der Mensch, meistens nicht vernünftig (oder nicht genügend schnell) auf Grund kontextueller Information auswählen können. Deshalb vermeiden General-Purpose-NLP-Systeme, die sich als Produkt bewähren müssen, eigentlich generell explizite semantische Repräsentationen. Der Charme eines Unterspezifikationsformalismus wie UDRT oder der Erweiterungsvariante FUDRT liegt darin, dass immer nur eine semantische Repräsentation des Satzes (mit klaren logischen Eigenschaften) vorliegt, die durch wachsende kontextuelle Constraints sukzessive zu einer Lesart verfeinert werden kann.

In der Diskursanalyse von Personal Translator werden aus Syntaxanalysen semantische Dependenzstrukturen konstruiert, die flache unterspezifizierte Diskursrepräsentationsstrukturen kodieren (vgl. Eberle (2002)). Anhand von Beispiel (9) können wir deren Gestalt und Nutzen für die Auflösung von Kennzeichnungen veranschaulichen:

(9)

- a) *Wenn eine Funktion graue Bildsegmente zeigt, macht sie von den Defaultregeln Gebrauch.*
- b) *Wenn sie graue Bildsegmente zeigt, macht eine Funktion von den Defaultregeln Gebrauch.*
- c) *Eine Funktion macht von den Defaultregeln Gebrauch, wenn sie graue Bildsegmente zeigt.*
- d) *Sie macht von den Defaultregeln Gebrauch, wenn eine Funktion graue Bildsegmente zeigt.*

Der semantische Filter verhält sich bezogen auf das Beispiel offensichtlich dann zufrieden stellend, wenn er für die Fälle a) - c) den Bezug des Pronomens *sie* auf *eine Funktion* erlaubt, bei d) jedoch nicht, oder diese Möglichkeit wenigstens mit einem starken Malus versieht. Allein daran, dass in d) das Pronomen kataphorisch verwendet wäre, kann es nicht liegen, in Anbetracht geglückter kataphorischer Bezüge (vgl. 1.b).

Wir versuchen, dieses Verhalten über FUDRS-Repräsentationen der Sätze und deren Spezifikationsmöglichkeiten zu erklären und daraus ein Filter-Kriterium zu gewinnen.

Zu (9.a) erhalten wir folgende FUDRS:

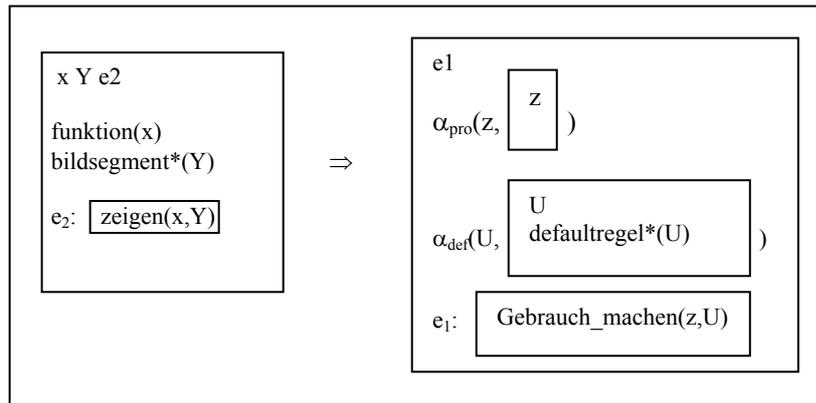
(9.a) FUDRS:

$$\left. \begin{array}{l} \text{vsubconj: } \underline{\text{wenn(zeigen)}_{e_2}} \\ \text{subj(n): } \underline{\text{sie}} \\ \text{comp(p(von))}: \underline{\text{den(defaultregeln)}_{\text{U}}} \end{array} \right\} \left. \begin{array}{l} \left. \begin{array}{l} \text{subj(n): } \underline{\text{eine(funktion)}_x} \\ \text{obj(n): } \underline{\text{graue(bildsegmente)}_y} \end{array} \right\} \& \text{OC}_1 \end{array} \right\} \& \text{OC}$$

Gebrauch machen_{e₁}

Dies ist wie folgt zu lesen: Der Satz ergibt eine Verbrepräsentation Gebrauch mache_{e₁} (die ein Ereignis e_1 einführt und vom Typ $(t/e/e)$ ist), auf die zwei Funktoren vom Typ $(t/(t/e))$ anzuwenden sind (die aus den subkategorierten Funktionen $\text{subj}(n)$ und $\text{comp}(p(\text{von}))$ stammen) und ein Funktor vom Typ (t/t) , aus dem Adjunkt vsubconj , wobei dieser Funktor sich wieder aus einer entsprechenden Satzrepräsentation zusammensetzt. Flach meint hier zweierlei. Erstens, lexikalische Unterscheidungen die beispielsweise *ein* als *spezifischen* Quantor oder als *generischen* oder als *unspezifisch-existentiellen* Quantor interpretieren sind (noch) nicht getroffen und in einem Ausdruck zusammengefasst (die unterstrichenen Ausdrücke stehen also für Funktionsterme, denen verschiedene Auswertungen als explizite Repräsentationen des jeweils geforderten Typs zugeordnet sind). Zweitens, die Reihenfolge der Applikation ist (noch) nicht festgelegt, bzw., genauer, ist bestimmt durch die Menge der jeweils zugeordneten *order constraints* (OC), die zu Beginn (bevor kontextuelle Einflüsse berücksichtigt werden) allerdings meist leer ist. Es werden in den FUDRSen die syntaktischen Funktionen der Rollen als Information aus der Syntax-Semantik-Schnittstelle festgehalten, einerseits, um den logischen Typ der Applikation zu charakterisieren ($\text{subj}(n)$ bezieht sich als $t/(t/e)$ -Funktors auf den Subjektsreferenten der Verbrepräsentation), andererseits um die durch den Typ gegebene semantische Rolle des Argumentreferenten nur unterspezifiziert zu beschreiben ($\text{subj}(n)$ als AGENS oder THEMA etc.).

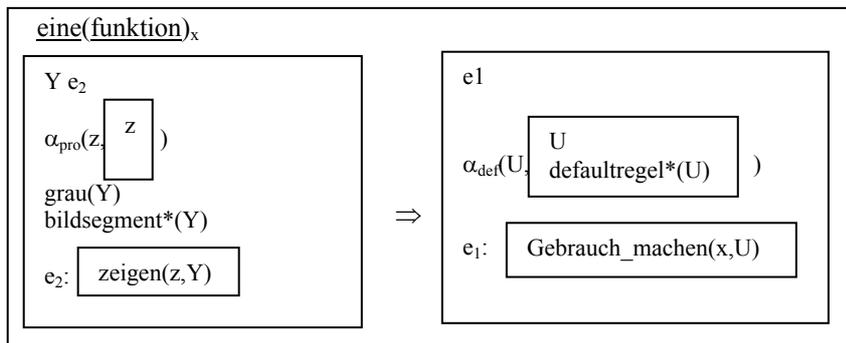
Aufgrund der Oberflächenordnung und des Typs der Modifikation nahe liegend ist im Fall (9.a), dass der vsubconj -Operator Skopus über den $\text{subj}(n)$ -Operator nimmt. Bei entsprechender weiterer Ausdifferenzierung der *order constraints*-Mengen bis zu jeweils eindeutigen Reihenfolgen, erhalten wir, nach Anwendung der Operatoren entsprechend der aufgefüllten OC und OC_1 (und kanonischer lexikalischer Interpretation), die folgende DRS, die insofern noch nicht ganz spezifisch ist, als die Präsuppositions/Resolutions-Trigger, die gekennzeichnet sind durch so genannte α -Konditionen, noch nicht ausgewertet sind:



Nach den DRT-Zugänglichkeitsbedingungen kann die Präsupposition zu z in der Restriktor-DRS gebunden werden, dh. z als x identifiziert, also *sie* auf *eine Funktion* bezogen werden.

Bei (9.b) erhalten wir dieselbe FUDRS wie bei (9.a), nur sind die Positionen von *sie* und *eine Funktion* in der Repräsentation vertauscht. Ein Bezug von *sie* auf *eine Funktion* ist in diesem Fall nur möglich, wenn das indefinite Subjekt Skopus über die wenn/dann-Repräsentation erhält. Eine FUDRS' die in der Weise spezifiziert ist, die gleichzeitig die anderen Funktoren mit engem Skopus anwendet und wie in (9.a) lexikalisch interpretiert, insbesondere die definiten Funktoren als referentiell, den indefiniten existentiell, aber die Entscheidung für *eine Funktion* noch offen lässt, hat die folgende Gestalt:

(9.b) FUDRS':



Auch hier kann z als x identifiziert werden, weil x als DRF der die Duplexbedingung einführenden DRS für z aus der Restriktor-Box dieser Bedingung zugänglich ist.

Für (9.c) erhalten wir dieselbe FUDRS wie bei b). Nur ist an der Oberfläche das Matrix-Subjekt vor der wenn-Subordination. Damit präferieren wir hier stärker noch als bei b) die Lösung mit weitem Skopus für das indefinite Subjekt mit der nämlichen Auflösungsmöglichkeit für das Pronomen. Bei d) liegt der Fall anders. Wir erhalten die folgende FUDRS:

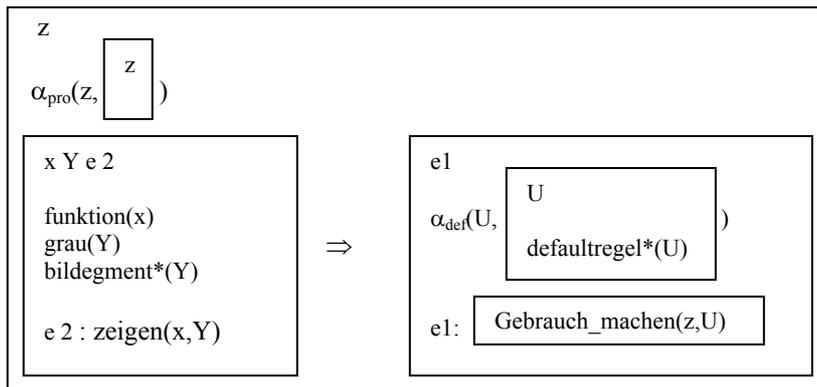
(9.d) FUDRS:

$$\left\{ \begin{array}{l} \text{subj}(n): \underline{\text{sie}}_z \\ \text{comp}(p(\text{von})): \underline{\text{den}(\text{defaultregeln})}_U \\ \text{vsubconj}: \underline{\text{wenn}(\text{zeigen}_{e2})} \end{array} \right\} \left\{ \begin{array}{l} \text{subj}(n): \underline{\text{eine}(\text{funktion})}_x \\ \text{obj}(n): \underline{\text{graue}(\text{bildsegmente})}_Y \end{array} \right\} \& \text{OC}_1$$

Gebrauch machen_{e1}

& OC

Wieder können wir davon ausgehen, dass ohne zusätzliche Annahmen OC und OC₁ zunächst leer sind. Das Pronomen *sie* in dieser Situation auf *eine Funktion* beziehen wollen, hieße, dem Matrix-Subjekt engen Skopus gegenüber dem wenn-Operator zu geben, obwohl es an der Oberfläche vor diesem Operator an der ersten Stelle, also an der kanonischen Stelle für das schon bekannte Thema steht (das als solches natürlich weiten Skopus haben muss). Deshalb stark präferiert oder überhaupt nur möglich scheint die Ausdifferenzierung zur folgenden DRS:



Hier kann z nur in der Haupt-DRS aufgelöst, d.h. nicht mit x identifiziert werden. Die Repräsentationen für b) und c) sind noch nicht völlig ausdifferenziert. Sie lassen, ne-

ben der existentiellen Interpretation, bei der der konditionale Zusammenhang für gerade **eine** in die Haupt-DRS eingeführte *Funktion* gefordert wird, auch die Lesart als generisches 'Gesetz' für *Funktionen* zu, die hier vielleicht angemessener ist (und auch noch eine spezifische Lesart). Wichtig ist: In jeder dieser Auswertungen des unspezifischen Quantors **ein** bleibt die Zugänglichkeit von x für z erhalten (im relevanten generischen Fall deshalb weil zwar eine neue Duplexbedingung entsteht, die alte aber in den Skopus dieser neuen Bedingung fällt). Für d) würden die alternativen Interpretationsmöglichkeiten für *eine Funktion* die fehlende Zugänglichkeit nicht aufheben!

Zu Hilfe genommen haben wir bei dieser Erklärung Annahmen über den Einfluss von Oberflächenanordnung auf die semantischen Skopusverhältnisse und über eine gewisse Rangfolge von Skopus tragenden Elementen bezogen auf ihre Modifikationscharakteristik. Beide Typen von Annahmen folgen aber gängigen Beobachtungen und Theoriebildungen aus der Literatur (vgl. Frey (1993), Pafel (1988)).

Wie funktioniert nun der entsprechende Filter? Falls bei der Bewertung eines Antezedent-Kandidaten festgestellt wird, dass der Bezug nur für bestimmte Skopusverhältnisse statthaft ist, wird aus der Information zur Oberflächenanordnung und einer vordefinierten Rangordnung Skopus tragender Elemente die Akzeptanz dieser Skopusanordnungen positiv oder negativ bewertet und, falls das Ergebnis nicht zu einem sofortigen Ausschließen der Bindungsmöglichkeit führt, mit den anderen Gewichten verrechnet.³ Ganz wesentlich ist in jedem Fall, wenn dieser Bezug schließlich gewählt wird, dann müssen die Ordnungsbedingungen der FUDRS auf die dann nur noch zugelassenen Fälle eingeschränkt werden. Denn eventuell folgende Referenzbildungen sind durch die aus der Wahl ableitbaren schärferen Zugänglichkeitsbedingungen beschränkt. Das bedeutet, Skopusverhältnisse und Referenzbildung müssen (partiell) zusammen entschieden werden!

Neben Quantifikationsstrukturen sind andere signifikante Strukturkonstellationen, die durch den Filter einer Untersuchung unterzogen werden, solche bei denen die Möglichkeit besteht, dass Antezedent-Kandidat und Anapher in unterschiedlichen Diskursmodi erscheinen (und dadurch Bezug ausgeschlossen ist), der Kandidat in einer modalen Einbettung zu einer reinen Hypothese oder zu einer deontischen Einstellung oder dergleichen, jedenfalls, das ist wichtig, ohne per Präsupposition nahe gelegter (de re-) Existenz, und das Pronomen außerhalb einer solchen Einbettung, wie in dem Beispiel (10). Dies bedeutet im Übrigen, dass sich solche Untersuchungen nur auf Antezedent-Kandidaten beziehen, die indefinit introduziert sind. Bei definiten Kennzeichnungen und Namen muss immer zumindest auch von der de re-Lesart ausgegangen werden und damit generell von Zugänglichkeit.

³ Das Systemhintergrundwissen zu Ordnungspräferenzen wird beispielsweise sagen, dass indefinite NPs weiten Skopus haben können gegenüber wenn/dann-Einbettung und bei Anordnung vor konditionalen Subordinationen in der Regel auch haben.

- (10) *Peter gesteht Irene seinen Traum. Er würde eine Traumfrau kennen lernen.*
a) *Sie würde lachen, fröhlich sein, jung sein.*
b) *Sie lacht.*

Wir deuten nur an, wie im Fall modaler Einbettungen vorgegangen wird: Wenn für den möglichen Antezedenten davon ausgegangen werden kann, dass er im Universum einer modal eingebetteten DRS eingeführt ist, muss als Minimalbedingung für die Anapher dasselbe gelten, wobei die in Frage stehenden Modaloperationen deontisch/modal miteinander kompatibel sein müssen. Erkennungszeichen sind natürlich typischerweise Tempusformen und Modalverben. So kann in (10.a) das Pronomen *sie* sich auf *eine Traumfrau* beziehen, weil es durch die Tempusform genauso in einen hypothetischen Kontext gestellt ist, wie die indefinite NP (und es wird, wegen der Modalparallelität eher so bezogen). Bei b) ist das nicht so, dort bezieht sich *sie* eindeutig auf *Irene*.

Auch in diesen Fällen gilt: Falls der gewählte Bezug die Applikationsbedingungen der FUDRS berührt, müssen diese in den entsprechenden order constraints passend spezifiziert werden.

4 Diskussion

Zur Evaluierung kann man die Koreferenzfälle im Text betrachten oder die Pronomeninterpretation als solche, was die pleonastische Verwendung einschließt und die Fälle von Pronomen-Akkommodierung. Unter dem Aspekt der Übersetzung ist in jedem Falle die zweite Perspektive vorzuziehen, weil die Pronomen mit und ohne Bezug richtig übersetzt werden müssen und weil Akkommodierung in dem für die Übersetzung typischen Einzelsatzmodus recht häufig ist.

Für Stichprobentexte unterschiedlicher Genres mit jeweils etwa 20 - 30 (rund 15 Wörter langen) Sätzen konnten wir Verbesserungen zwischen 4 % und 10 %, in Einzelfällen auch darüber, verzeichnen. Da Morphologie und Part of Speech-Ambiguität für Systeme mit großen Lexica und Analysekapazitäten wie *Personal Translator* nur ein marginales Problem sind, variiert das Erkennen und morphologisch korrekte Beschreiben der Pronomen im Text nahe 100 % (abhängig von den Ambiguitätsspezifika der betrachteten Sprache). Will heißen, unter der eingenommenen Betrachtungsweise beschreiben diese Werte sowohl Präzision als auch Recall.

Woher rührt die große Varianz? Es wurden deutsche, englische und französische Texte sehr unterschiedlichen Typs getestet, von Ausschnitten aus technischen Dokumentationen bis zu Dialogpassagen aus Romanen mit direkter und indirekter Rede und

Sammlungen von Einzelsätzen, mit großen Unterschieden was Anzahl und formale Unterscheidbarkeit der Diskursreferenten betrifft. Besonders dann natürlich, wenn mehrere Diskursreferenten auseinander zu halten sind, die morphologisch ähnlich sind, oder wenn es mehrere Diskursebenen gibt, wirkt sich der semantische Filter sehr positiv aus. Im Schnitt, das bestätigen auch Tests zur Veränderung des Übersetzungsverhaltens zwischen den genannten Sprachen die an Testsuites mit jeweils mehreren 10.000 Sätzen durchgeführt wurden, werden durch den Einbezug des Filters etwa 6 % der Sätze mit Pronomenverwendung besser übersetzt.

Warum werden keine (noch) höheren Werte erzielt? Der syntaktische Filter arbeitet nur korrekt, wenn die syntaktische Analyse des Satzes weitgehend korrekt ist und diese Voraussetzung ist gerade bei längeren Sätzen über 10 oder 15 Wörtern immer weniger gewährleistet. Weil die semantische Analyse sich auf die syntaktische bezieht, verfälschen solche Fehler auch die Daten für die semantische Untersuchung. Eine weitere Quelle für Irrtümer bei der semantischen Analyse sind unvollständige oder falsche Hintergrund-Informationen, weil die Ressourcen, vor allem Lexika, nicht für jedes Phänomen oder Wort gleich vollständig und sauber kodiert sind. Und es sind durch Grice'sche Maximen angestoßene Reinterpretationen, vor allem Metonymie (wobei das System eine Anzahl regulärer Fälle durchaus behandeln kann).

Die so genannten wissensarmen (knowledge poor) Ansätze vermeiden solche Probleme natürlich (vgl. u.a. Kennedy/Boguraev (1996), Baldwin (1997)). In der Regel beschränken sich entsprechende Vorschläge auf morphologische Filter und höchstens schwach analysierte, meist PoS-getaggte Texte und Bewertungen signifikanter struktureller Pattern, eventuell über Corpora gelernt und trainiert. Dabei spielt häufig Semantik eine Rolle, zuallermeist über (gelernte) Kollokationen und Selektionsbeschränkungen. Die Literatur berichtet für einige solche Ansätze beeindruckend hohe Präzisionswerte (von über 90 %). Nur sind diese in beinahe allen Fällen erkauft durch Bezug auf eine sehr eingeschränkte Domäne oder äußerst vorsichtiges Auflösen, in der Konsequenz also einen niedrigen Recall (bei 90 % Präzision typischerweise 60 %, vgl. den Überblick in Mitkov (1999)).

Wenn Anaphernresolution Modul eines regelbasierten NLP-Systems ist, das ohnehin syntaktische (und semantische) Repräsentationen für die Zwecke seiner zentralen Funktionalität aufbaut und entsprechende Ressourcen vorsieht, macht es auf Grund dieser Resultate keinen Sinn, auf diese Informationen zu verzichten und für die Resolution einen wissensarmen Ansatz zu fahren. Es macht aber sehr wohl Sinn, und wird die weitere Forschung und Entwicklung ein Stück weit bestimmen, Heuristiken aus solchen Verfahren zur Fehlerkorrektur und Vervollständigung des regelbasiert syntaktisch-semantischen Ansatzes (auch durch statistisch fundiertes Lernen semantischer Zusammenhänge aus Corpora) einzusetzen, um noch größere Verlässlichkeit zu erzwingen.

Literatur

- Baldwin, B. (1997): CogNIAC: high precision coreference with limited knowledge and linguistic resources. In: Proceedings of the ACL'97/EACL'97 workshop on Operational factors in practical, robust anaphora resolution, Madrid, Spain, 38-45.
- Bosch, P./Geurts, B. (1990): Processing definite NPs. *Rivista di Linguistica*, 2, 177-199.
- Carbonell, J. G./Brown, R. D. (1988): Anaphora resolution: a multi-strategy approach. In: Proceedings of the 12. International Conference on Computational Linguistics (COLING'88), Budapest, Hungary, 96-101.
- Carter, D. M. (1986): A shallow processing approach to anaphor resolution. PhD thesis, University of Cambridge, Cambridge.
- Chomsky, N. (1981): Lectures on Government and Binding. Dordrecht: Foris.
- Connolly, D./Burger, J. D./Day, D. S. (1994): A Machine learning approach to anaphoric reference. In: Proceedings of the International Conference New Methods in Language Processing, Manchester, United Kingdom, 255-261.
- Dagan, I./Itai, A. (1990): Automatic Processing of Large Corpora for the Resolution of Anaphora References. In: Proceedings of Coling90.
- Eberle, K. (1997): Flat underspecified representation and its meaning for a fragment of German. *Arbeitspapiere des Sonderforschungsbereichs 340 Sprachtheoretische Grundlagen für die Computerlinguistik* 120, Stuttgart: Universität Stuttgart.
- Eberle, K. (2002): Tense and Aspect Information in a FUDR-based German French Machine Translation System. In: Kamp, H./Reyle, U. (eds.), *How we say WHEN it happens*. Tübingen: Niemeyer, 2002. *Ling. Arbeiten*, Band 455, 97-148.
- Frey, W. (1993): Syntaktische Bedingungen für die semantische Interpretation. Akademie Verlag, Berlin, *Studia Grammatica* Bd. XXXV edition.
- Guenther, F./Lehmann, H. (1983): Rules for Pronominalization. In: Proceedings of the First European ACL Meeting, 144-151.
- Heim, I. (1983): File Change Semantics and the Familiarity Theory of Definiteness. In: Bäuerle, R./Schwarze, R./von Stechow, A. (eds.), *Meaning, Use and Interpretation of Language*. Berlin: de Gruyter.
- Kamp, H. (1981): A Theory of Truth and Semantic Representation. In: Groenendijk, J./Janssen, T./Stokhof, M. (eds.), *Formal Methods in the Study of Language*. Mathematical Centre Tract, Amsterdam.
- Kamp, H./Reyle, U. (1993): *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.
- Kennedy, C./Boguraev, B. (1996): Anaphora for everyone: pronominal anaphora resolution without a parser. In: Proceedings of Coling96, Copenhagen, Denmark, 1996, 113-118.
- Lappin, S./McCord, M. (1990): Anaphora Resolution in Slot Grammar. *Computational Linguistics*, 16.
- Lappin, S./Leass, H. (1994): An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4), 535-561.
- Lasnik, H. (1976): Remarks on Coreference. *Linguistic Analysis*, 2(1).
- Leass, H./Schwall, U. (1991): *An Anaphora Resolution Procedure for Machine Translation*. IWBS Report 172, IBM Deutschland, Heidelberg: Wissenschaftszentrum Heidelberg.

- McCord, M. (1991): The Slot Grammar System. In: Wedekind, J./Rohrer, C. (eds.), *Unification in Grammar*. MIT-Press.
- Mitkov, R. (1998b): Robust pronoun resolution with limited knowledge. In: *Proceedings of the 18th International Conference on Computational Linguistics (Coling-98)*, Montreal, Canada.
- Mitkov, R. (1999): *Anaphora Resolution: The State of the Art*. working paper, University of Wolverhampton.
- Nakaiwa, H./Ikehara, S. (1995): Intrasentential resolution of Japanese zero pronouns in a Machine Translation system using semantic and pragmatic constraints. In: *Proceedings of the International Conference on Theoretical and Methodological Issues in Machine Translation (TMI'95)*, Leuven, Belgium, 96-105.
- Pafel, J. (1988): *Die Parameter des relativen Quantorenskopus im Deutschen*. LILOG Report 48, IBM Deutschland, TK LILOG, Stuttgart.
- Pollard, C./Sag, I. (1994): *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.
- Reinhardt, T. (1983): *Anaphora and semantic interpretation*. Chicago: University of Chicago Press.
- Reyle, U. (1993): Dealing with ambiguities by underspecification: Construction, representation, and deduction. *Journal of Semantics*, 10(2):123-179.
- van der Sandt, R. (1992): Presupposition Projection as Anaphora Resolution. *Journal of Semantics*, 9(4).