gefördert vom Bundesministerium für Forschung und Technologie

Verbmobil
Verbundvorhaben

Semantik-orientierter rekursiver Transfer in HPSG am Beispiel des Referenzdialogs

Michael Dorna
Kurt Eberle
Martin Emele
C.J. Rupp

Universität Stuttgart

Report Nr. 39
Oktober 1994

Oktober 1994

Michael Dorna
Kurt Eberle
Martin Emele
C.J. Rupp

Institut für maschinelle Sprachverarbeitung – Computerlinguistik
Universität Stuttgart
Azenbergstraße 12
D–70174 Stuttgart

Tel.: (0711) 121 - 1363
Fax: (0711) 121 - 1366
e-mail: {dorna,eberle,emele,rupp}@ims.uni-stuttgart.de

Gehört zum Antragsabschnitt: 12.1., 12.2., 12.5.

Die vorliegende Arbeit wurde im Rahmen des Verbundvorhabens Verbmobil vom Bundesministerium für Forschung und Technologie (BMFT) unter dem Förderkennzeichen 01 IV 101 G gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei den Autoren.

Inhaltsverzeichnis

Abstract

Wir machen einen Transfervorschlag für HPSG-Strukturen, der bei der Übersetzung der semantischen Repräsentation die syntaktische Struktur zusammen mit dem HPSG-spezifischen *Semantischen Prinzip* (Pollard/Sag(1994)) als rekursionssteuerndes Moment benutzt. Der Transfer übernimmt bei dieser Architektur eine Mediatorrolle zwischen dem Analyseergebnis, das eine quellsprachenspezifische semantische Repräsentation enthält, und dem Generatorinput, der eine zielsprachenspezifische semantische Repräsentation enthält. Das heißt u.a., daß die Transferkomponente Übersetzungsschwierigkeiten feststellt, strukturelle Veränderungen vornehmen und, bei Bedarf, zur Auflösung von *translation mismatches* (Kameyama et al.(1991)), eine vertiefte semantische Analyse der quellsprachlichen Äußerung anstoßen und deren Ergebnisse verwenden kann.

Als mittelfristiges Ziel soll die Generierung zur Herstellung von Übersetzungsvarianten den Transferoutput an die Transferkomponente zurückgeben können mit Maßgaben, die es dem Transfer erlauben, gezielt alternative zielsprachspezifische Repräsentationen zu erzeugen. Unter dem Aspekt dieser kommunikativen Kompetenz übernimmt der Transfer die Rolle eines *negociators* (Kay et al.(1994)).

1 Vorbemerkung

In Verbmobil wird ein modulares, semantik-orientiertes Übersetzungskonzept verfolgt. Quellsprachlichen Äußerungen werden semantische Repräsentationen zugeordnet. Zielsprachliche Äußerungen werden aus semantischen Repräsentationen generiert, unter Zuhilfenahme spezifischer nicht-semantischer Information, wie pragmatischer Information, Fokus- und Topikalisierungsinformation. Unter der Prämisse, daß die Generierung zur Erfüllung ihrer Aufgabe in der Regel keine weiteren Anfragen an die quellsprachliche Analyse stellt, liegt die eigentliche Übersetzungsleistung bei dem Modul, der die semantische Repräsentation, aus der generiert werden soll, herstellt: beim Transfermodul.

Die semantik-orientierte Übersetzung kennt im wesentlichen zwei Ausrichtungen: Quell- und zielsprachlicher Ausdruck referieren auf dieselbe semantische Repräsentation. In diesem Fall dient die semantische Repräsentation als Interlingua bezogen auf das betrachtete Sprachpaar (oder die betrachteten Sprachpaare). Ein neuerer auf die HPSG bezogener Ansatz findet sich in Kuhn/Heid(1994), der sogenannte shared-content-Ansatz. Alternativ referieren quell- und zielsprachlicher Ausdruck auf verschiedene semantische Repräsentationen, auf Repräsentationen, die bezogen auf ihren propositionalen Gehalt (vor dem Hintergrund eines äquivalenten Kontexts) gleich, oder zumindest einander sehr ähnlich sind, die in ihrer Form aber voneinander abweichen (können).

In Verbmobil erfolgt die semantische Analyse im Rahmen der *Diskursrepräsentationstheorie* (DRT). Semantische Repräsentationen sind (partielle) *Diskursrepräsentationsstrukturen* (DRSen) mit ausgezeichnetem *Diskursreferent* (DRF). Wir geben das hier in der folgenden Notation wieder: $(\lambda \ldots) < x, D_x >$, wobei x der ausgezeichnete Referent ist, der in der DRS D_x eingeführt oder thematisiert ist[1].

Die Semantikkonstruktion in Verbmobil[2] stellt aus guten Gründen, zur Vermeidung unnötigen inferentiellen Aufwands an der falschen Stelle und zur weitestgehenden Vermeidung des Mitführens von Disjunktionen, für quellsprachliche Äußerungen nur schwach spezifizierte DRSen zur Verfügung. So werden Ambiguitäten nur dann aufgelöst oder disjunktiv repräsentiert, wenn sie mit strukturell divergenten Repräsentationsformen assoziiert sind, nicht aber, wenn sie in einer einheitlichen, wenn auch wenig spezifischen Weise repräsentiert werden können. Präpositionen beispielsweise sind häufig mehrdeutig, man vergleiche etwa *ich komme dann mit dem Auto* versus *ich komme dann mit meinem*

[1] Zur verwendeten Notation vergleiche man etwa Eberle(1991).
[2] Vgl. Bos et al.(1994) und Bos/McGlashan(1994).

Angestellten K., wobei im ersten Fall *mit* das Ereignis mit einer *Instrument*-Rolle versieht, im zweiten Fall mit einer *Commitativ*-Rolle. Präpositionen sind aber häufig mehrdeutig nur in einem nicht-strukturellen Sinn, dh. ihre verschiedenen Bedeutungen sind alle vom selben semantischen Typ und können deshalb in einer gemeinsamen Form untergebracht werden, *mit* etwa in $\lambda < x, D_x > \lambda < e, D_e >< e, \boxed{\texttt{mit(e,x)}} \& D_x \& D_e >$.

Eine solche Repräsentationsphilosophie ist nur schwer mit einem Interlingua-Ansatz zu verbinden. Das Englische beispielsweise macht zwar auch von mehrdeutigen Sprachmitteln Gebrauch, aber, wie das Beispiel zeigt, mit einer vom Deutschen verschiedenen Strukturierung des ontologischen Bereichs. *Mit dem Auto* würde man durch **by** *car* übersetzen, *mit dem Angestellten* jedoch durch **with** *the employee*, wobei die Präposition *by* im Deutschen wiederum zwar (instrumentales) *mit* ergeben kann, aber auch *durch* oder *von* etc. Eine Interlingua hätte in der Granularität ihrer denotierenden Ausdrücke wenigstens so fein strukturiert zu sein, wie jede der im Übersetzungssystem berücksichtigten natürlichen Sprachen. Das heißt, der Interlingua/shared-content-Ansatz kann auf Konditionen wie *mit(e,x), by(e,y), durch(e,z)* etc. verzichten, sie sind zu wenig spezifisch angesichts des Wortauswahlproblems bei der Generierung. Er benötigt stattdessen die Domäne feiner strukturierende Relationen wie *commitativ(e,x)* oder *causa_finalis(e,x)* oder *causa_materialis(e,x)* etc.

Der skizzierte Repräsentationstyp fügt sich aber offensichtlich sehr gut in das zweite Übersetzungskonzept mit Source- und Target-DRSen deren Prädikate nach dem Granularitätsgrad der jeweiligen Bezugssprache spezifisch gewählt sind. Somit wurde der Shared-content-Ansatz, mit DRSen als Interlingua-Ausdrücke, zwar diskutiert, aber für Verbmobil verworfen. Man beachte, daß der zweite Ansatz durchaus Shared-content-Beschreibungen für spezifische Phänomene erlaubt, etwa für die Analyse der Tempus-Information.

Eine ökonomische Variante der Übersetzung vermöge sprachspezifischer DRSen ist, unmittelbar die Target-DRS zu erzeugen und die Erstellung der Source-DRS auszulassen. Dabei werden die Source-Lexikoneinträge direkt mit Target-Semantik assoziiert und die Semantikkonstruktion produziert über der Konstituentenstruktur des Quellausdrucks mithilfe solcher Lexikoneinträge direkt die Target-Semantik zur Äußerung aus der dann generiert wird. Diese Variante (vgl. C.J.Rupp et al.(1994)) wurde verworfen aufgrund vermuteter inferentieller Unzulänglichkeit, erwarteter Divergenzen im Konstruktionsprozess und wegen des eingeforderten Bedarfs nach getrennten Repräsentationen für Source- und Target-Äußerung im größeren Text- oder Dialog-Rahmen (verschiedene Verweismöglichkeiten etc.).

Damit gehen wir von folgendem aus: Die Source-DRS führt eine Menge von Diskursreferenten ein und Konditionen (Relationen) über diese DRFs, deren Strukturierungseigenschaften am Lexem-Sortiment der Quellsprache orientiert sind (und entsprechend für Sub-DRSen der Source-DRS). Der Transfermodul ordnet der Source-DRS eine Target-DRS zu, für die das nämliche bezogen auf die Zielsprache gilt.

DRSen sind in der Regel zu schwach strukturiert, um sinnvoll rekursiven Transfer zu gestatten. So ist beispielsweise der DRS-Konditionenliste nicht anzusehen, welche Konditionen derselben Satzkonstituente entstammen etc. Auf genau solche Zusammenhänge wird man aber rekurrieren wollen, um Transferregeln anbieten zu können, die Konditionen nicht nur eins zu eins abbilden, sondern auch bestimmte Konditionenbündel geschickt zusammenfassen.[3]. Deshalb wird die Transfer-Rekursion nicht an der DRS als solcher festgemacht, sondern an der Daughters-Struktur des HPSG-Source-Signs. Mit den Töchterknoten des Source-Signs sind vermöge des HPSG-spezifischen Semantischen Prinzips (*Semantics Principle*) syntaktisch sinnvoll motivierte Teil-Semantiken assoziiert. Die Formulierung des Transfers für ganze Signs, nicht bloß für inkorporierte DRSen hat auch den Vorteil, daß bei der Formulierung der Transferregeln nicht nur semantische Bedingungen abgeprüft werden können, sondern auch Anfragen nach der syntaktischen Realisierung möglich sind. Gleichzeitig ist pragmatische bzw. Sprechakt-Information zugänglich, kann verwendet, und auch an das Target-Sign

[3] Shake-and-bake-Transfer auf Repräsentationen, Whitelock(1992), ist attraktiv, weil er genau solche Bündelungen erlaubt, der Nachteil ist aber, daß die Menge der prinzipiell möglichen Bündelungen exponentiell ist in der Menge der Konditionen. Ohne adäquate Einschränkungen, wie hier über die gegebene Satzstruktur, bleiben solche Verfahren ineffizient.

weitergereicht werden.

Bild 1 zeigt die Funktionalität des Transfer-Moduls in schematisierter Notation für das Beispiel (1). Der Input ist ein vollspezifiziertes Äußerungs-Sign, wobei in Bild 1 nur die für uns im folgenden relevanten Pfade notiert sind.[4] Der Output ist ein unterspezifiziertes Sign. In jedem Fall jedoch ist der *drs*-Pfad mit einer vollspezifizierten DRS bewertet und die Pragmatik-Information übergeben. Optional können andere Informationen notiert sein, und optional mag auch die Pragmatik-Information sich in Details ändern. Die Übersetzungsrelation *tau* führt, bezogen auf die semantische Repräsentation, eine Variablenumbenennung durch (abhängig davon welche Struktur in Verbmobil ein DRF hat, mag die Umbenennung auch die Identität sein), und ersetzt Source-sprachspezifische Konditionen durch Target-sprachspezifische. Die Übersetzungsrelation *tau* ist durch C, den Kontext, restringiert. Die Töchterknoten MORGEN und ICH_TREFFE_SIE im Source-Sign deuten an, wie in der Rekursion die Gesamt-Semantik aufgebrochen und abgearbeitet wird. Bild 2 zeigt das allgemeine Format einer

(1) *Ich treffe Sie (dann) morgen*

phon: <ich treffe Sie morgen>
prag: P
drs:
n t i u e
e: $\boxed{\text{treffen(i,u)}}$
theme1(e)=i
theme2(e)=u
e $\oslash$ t
next_day(n)=t

<n,sprechzeit>
<i,sprecher>
<u,hörer>

dtrs: [adj: MORGEN, head: ICH_TREFFE_SIE]

⇐ Input: vollspezifiziertes Sign

$\Downarrow \mathrm{tau}_{|C}$

prag: P
drs:
n_t t_t i_t u_t e_t
e_t: $\boxed{\text{meet}(i_t,u_t)}$
theme1(e_t)=i_t
theme2(e_t)=u_t
$e_t \oslash t_t$
next_day(n_t)=t_t

<n_t,sprechzeit>
<i_t,sprecher>
<u_t,hörer>

⇐ Output: unterspezifiziertes Sign mit vollspezifizierter Target-DRS

Abbildung 1: Die Funktionalität

tau-Regel: *tau* überführt das Source-Sign S in das unterspezifizierte Target-Sign T unter der Voraussetzung des Kontexts C, falls die spezifizierten *"Conditions"* erfüllt werden können. Im nächsten

[4] Es ist klar, daß *drs* kein Sign-Feature ist, sondern erst tiefer eingebettet erscheint. Aus Gründen der leichteren Lesbarkeit ist aber hier wie auch an anderer Stelle vereinfacht worden. Zu einer implementierungsnäheren Notation vergleiche man den Abschnitt *Implementierung*.

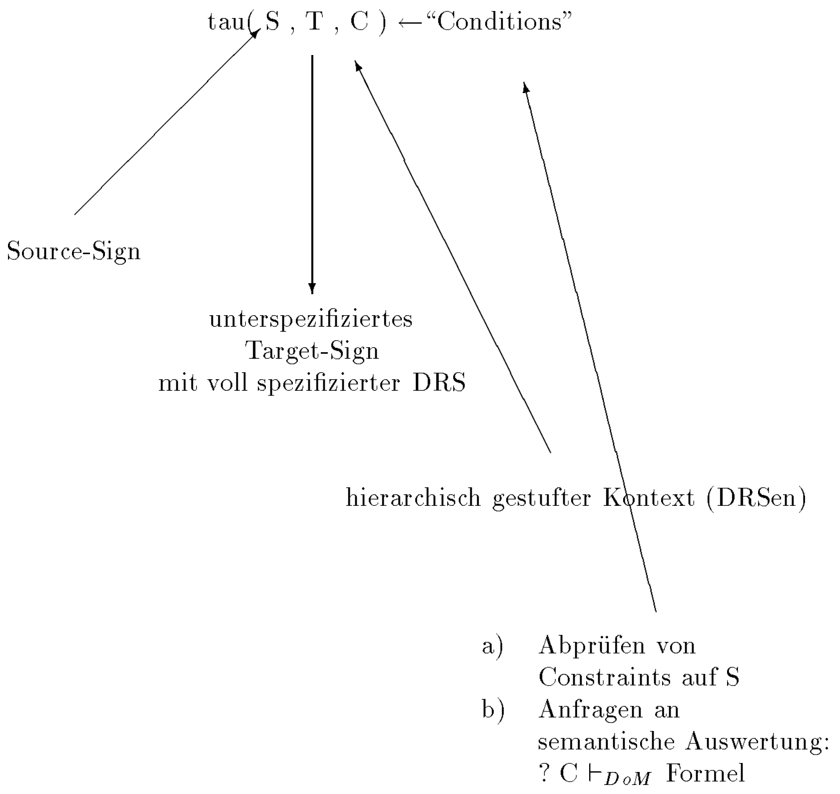

Abbildung 2: Das Format der *tau*-Regel

Abschnitt werden wir für das Beispiel (1) die Semantikkonstruktion skizzieren, um zu illustrieren an welchen Knoten der Töchterstruktur des HPSG-Signs welche semantische Information erhältlich ist. Daraus können wir dann eine Transfer-Strategie ableiten: Wir durchwandern das HPSG-Sign von oben nach unten entlang der semantischen Projektionslinie bis zum semantischen Kopf des Signs. Für diesen Kopf existiert eine *tau*-Regel im Transfer-Lexikon, die rekursiv Übersetzungen für die Argumente dieses Kopfes fordert und diese Übersetzungen in die Übersetzung des Kopfes inkorporiert. Diese Argumente sind Knoten in der Töchterstruktur des Gesamt-Signs. Auf sie kann demnach in gleicher Weise die eben beschriebene Transfer-Strategie angewendet werden. Der Prozess terminiert bei Erreichen von saturierten Köpfen.

In Abschnitt 3 werden wir exemplarisch Regeln diskutieren, die gestatten ein Sign entlang der semantischen Projektionslinie(n) zu traversieren und das Semantische Prinzip für das Target-Sign zu garantieren. In Abschnitt 4 werden die Transfer-Lexikoneinträge für Beispiel (1) skizziert. In Abschnitt 5 wird ein etwas komplizierteres Beispiel aus dem Referenzdialog diskutiert, in Abschnitt 6 wird auf Besonderheiten bei eingebetteten semantischen Strukturen hingewiesen und in Abschnitt 7 wird auf die Umschrift bei der tatsächlichen Implementierung eingegangen.

Betont werden muß, daß die vorgeschlagene Architektur dem Komplex Transfer/Generierung hohe Freiheiten einräumt: Erstens, es stehen keine Transferbedingungen im Quellsprachen-Lexikon. Sie sind unabhängig in einem eigenen Transferlexikon spezifiziert. Wesentlich ist für die Absprache mit dem Lexikon-Teilprojekt nur, in Kooperation mit der Semantikkonstruktion, die Namen der semantischen Prädikate und die entsprechende Zugriffsfunktion festzulegen (vgl. dazu auch den Abschnitt 7 zur Implementierung). Zweitens, der Transfer beschreibt die Targetsemantik indirekt durch Verwendung von semantischen Macros (Abschnitt 7). Natürlich sollen in dieser Weise, durch

Verwendung derselben Typen von Macros, die Konstruktionsprinzipien der Source-Seite zugleich die Semantikkonstruktion der Target-Seite definieren. Durch eine je verschiedene Definition der Macros kann aber optional für die Targetseite ein ganz anderes Semantikformat als für die Sourceseite verwendet werden (beispielsweise die *Minimal Recursive Semantics* aus Copestake et al.(1994)). Entscheidend und ein wichtiger Freiheitsgrad ist, daß auch bei Wahl eines anderen Typs von Targetsemantik, die Transferarchitektur und die verwendeten Transferregeln gleich bleiben. Drittens, das vom Transfer angebotene Regelpaket wie es im folgenden skizziert wird, zielt auf die Erstellung einer zielsprachennahen Repräsentation. Es ist der Generierung aber unbenommen das Angebot sprachnaher DRS-Konditionen zu ignorieren und allein aus dem konzeptuellen Wissen (das hierarchisch entsprechend der DRS-Strukturierung angeordnet ist) zu generieren. Zu diesem Zweck kann die Generierung mittelfristig durch Kommunikation mit dem Transfer eine zusätzliche Ausdifferenzierung konzeptueller Information auf der Basis der Source- und der Target-DRS (mit aufeinanderbezogenen DRFs) über dem Domänenmodell anstoßen. Viertens, die Architektur läßt es prinzipiell zu alternativeTransfer-Regelpakete zu formulieren und zu benutzen, die auf eine rein konzeptuelle Darstellung zielen (wenigstens dann, wenn klar ist wie eine solche Darstellung im Einzelfall aussieht). Dies ist zwar für Verbmobil nicht angestrebt, kann aber zu Evaluationszwecken für eine Menge einschlägiger Testsätze erfolgen. In jedem Fall erlaubt, schließlich, das Mitführen des Kontexts in der tau-Regel, semantische Information auch über den lokal durch das Semantische Prinzip verfügbaren Teil hinaus, an jeder Stelle an der ein Übersetzungsproblem sichtbar wird, abfragen zu können.

2 Semantikkonstruktion

Bild 3, die um Stellungsphänomene vereinfachte Töchterstruktur zu Beispiel (1) zeigt, wie (1) syntaktisch-semantisch analysiert ist: Töchter des Gesamt-Signs sind die durch *morgen*, realisierte Adjunkt-Tochter und die durch *ich treffe Sie* realisierte Head-Tochter. Letzteres Sign wird analysiert in eine Head-Complement-Struktur, bestehend aus dem durch *treffe Sie* realisierten Head und dem durch *ich* realisierten Komplement. Die *treffe Sie*-Struktur schließlich hat als Kopf das Word-Sign zu *treffe*, das sich im Lexikon findet und als Komplement das Sign zu *Sie*, das ebenfalls, wie auch das Sign zu *ich*, lexikalisch ist. Die Semantik des Lexikoneintrags zu *treffe* erwartet Beschreibungen der thematischen Rollen *theme1* und *theme2* und inkorporiert diese in den weiterzureichenden Semantikbeitrag. Diese Beschreibungen werden dem Subkategorisierungsrahmen entnommen, dabei wird auch sichergestellt, daß die Beiträge der verschiedenen grammatischen Funktionen richtig auf die verschiedenen thematischen Rollen verteilt werden. Vorausgesetzt wird in dem vorgeschlagenen Ansatz, daß der Subkategorisierungsrahmen eine Liste von Signs, nicht bloß von *synsem*- oder *syn*-Strukturen enthält. Die Notation $< x, D_x >$ meint, wie schon im letzten Abschnitt beschrieben, eine DRS, D_x, mit ausgezeichnetem DRF, x. x ist bei DRSen zu VPs in der Regel der Ereignis-Referent, bei DRSen zu NPs ist es in der Regel der vom Kopf eingeführte Referent. Die Applikation der Head-Semantik auf die Komplement-Semantik, realisiert durch die Unifikation 2, die der syntaktischen Abarbeitung des Subkategorisierungsrahmens vermöge der Unifikation 1 folgt, ergibt die Struktur 6 mit instantiiertem D_y. Dem Semantischen Prinzip der HPSG folgend wird diese Struktur zum nächst höheren Knoten gereicht. Wir sehen dort, dh. bei *treffe Sie*, 6 tatsächlich auch mit der entsprechenden Instantiierung. Der Beitrag der Komplement-Semantik ist in diesem Fall die Bereitstellung eines DRFs für den Hörer ohne weitere DRS-Kondition. Daß der eingeführte DRF bei Interpretation in einem Modell auch tatsächlich auf den Hörer gebunden wird, garantiert der entsprechende *externe Anker*. Im nächsten Schritt wird nun ganz analog die Sprecher-Semantik durch die Unifikationen 4, bzw. 3 mit der erwarteten *theme1*-Beschreibung identifiziert und das Ergebnis zum Mutterknoten gereicht: 6 in der *ich treffe Sie*-Struktur. Die semantische Projektionslinie bricht an dieser Stelle ab, deshalb, weil die *ich treffe Sie*-Struktur semantisch Argument ist bezogen auf *morgen*. In der Adjunct-Head-Struktur ist das Adjunkt der semantische Kopf, der die Semantik des Heads als Argument nimmt und das Ergebnis zum Mutterknoten reicht. Durch die Unfikationen 6, bzw. 5, erhalten wir das Ergebnis 7, das die Semantik des Adjunkts und des

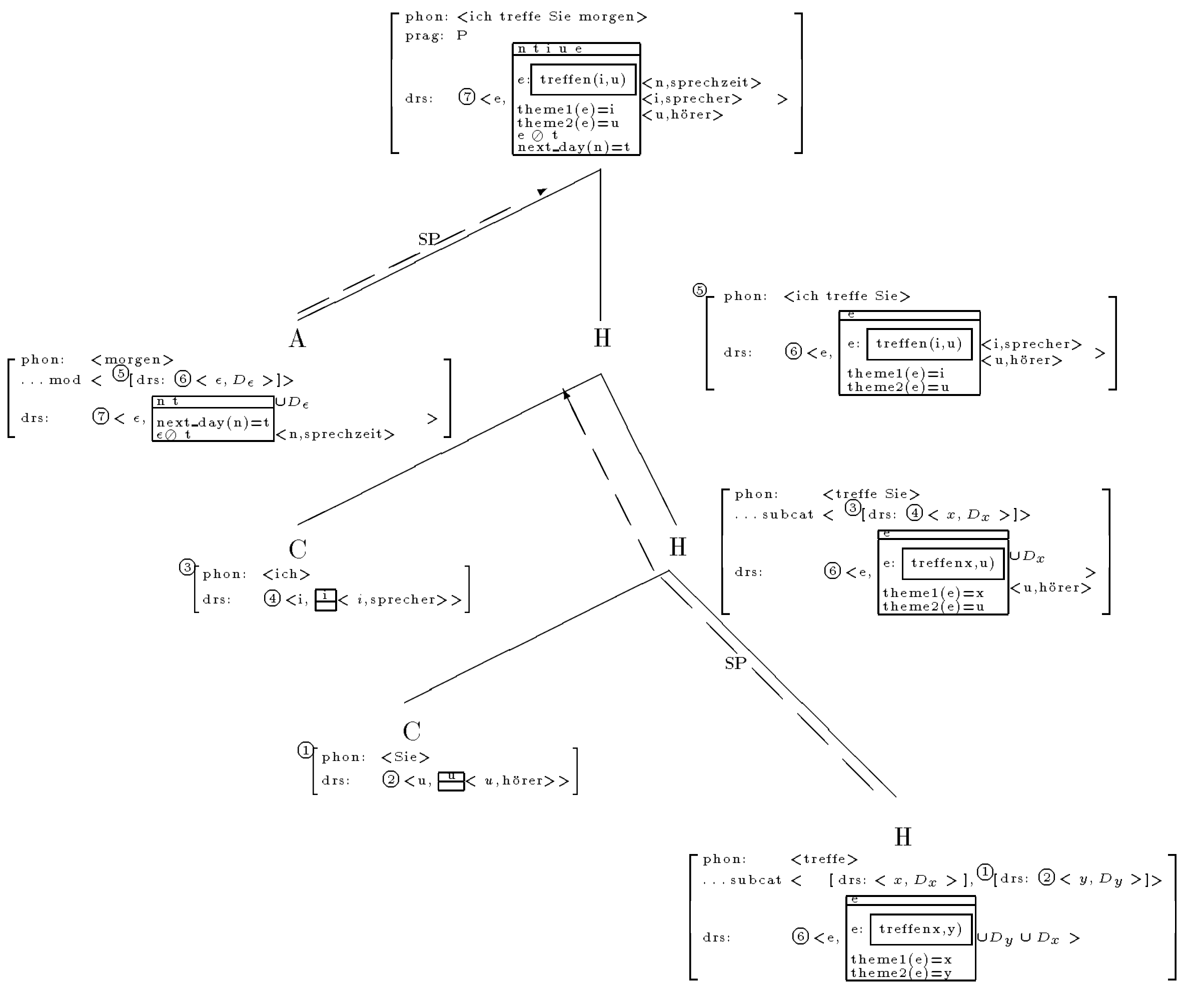

Abbildung 3: Die Töchterstruktur zu Beispiel (1)

Gesamt-Signs ist. Es ist wesentlich, sich vor Augen zu halten, daß entlang der Projektionslinien, nach durchgeführter Semantikkonstruktion, dieselbe Information steht. Beim Verbknoten also etwa das vollinstantiierte 6.

3 Perkolationsregeln für den Transfer

Es genügt, für den semantischen Kopf einer Äußerung die Target-DRS zu bestimmen, wegen des Semantischen Prinzips ist diese DRS auch die richtige Target-DRS für die gesamte Äußerung. Entsprechendes gilt für die Argumente des semantischen Kopfs. Wir benötigen demnach ein Sortiment von Perkolationsregeln das für jede in HPSG mögliche Konstellation von Töchterknoten die Target-Semantik des Signs bestimmt als die Target-Semantik der Tochter, die semantisch der Kopf ist. Im folgenden zur Illustration nur die beiden grundlegenden Fälle: *Adjunkt-Head-Struktur* und *Head-Complement-Struktur*:

- $\boxed{\text{Adjunct-Head-Struktur}}$

$$\text{tau}(S_{\left[\begin{array}{ll}\text{prag: P} & \\ \text{drs:} & \left[\begin{array}{ll}\text{adj:} & \text{A}\\ \text{head:} & \text{H}\end{array}\right]\end{array}\right]}, T_{\left[\begin{array}{ll}\text{prag: P} & \\ \text{drs:} & \text{TDRS}\end{array}\right]}, C)$$
$$\leftarrow \text{A}_{|\text{prag:P}}\ \&\ \text{tau}(A, TA_{[\text{drs: TDRS}]}, C).$$

- $\boxed{\text{Head-Comp-Struktur}}$

$$\text{tau}(S_{\left[\begin{array}{ll}\text{prag: P} & \\ \text{drs:} & \left[\begin{array}{ll}\text{head:} & \text{H}\\ \text{comp:} & \text{C}\end{array}\right]\end{array}\right]}, T_{\left[\begin{array}{ll}\text{prag: P} & \\ \text{drs:} & \text{TDRS}\end{array}\right]}, C)$$
$$\leftarrow \text{H}_{|\text{prag:P}}\ \&\ \text{tau}(H, TH_{[\text{drs: TDRS}]}, C).$$

Man liest dabei die erste Regel wiefolgt: bei gegebenem Kontext C wird das Sign S, dessen Struktur die als Index zu S gegebenen Anforderungen erfüllen muß, in das unterspezifizierte Sign T übersetzt, mit Feature-Werten wie im T-Index, wenn die Adjunkt-Struktur A aus S, erweitert um die nach unten gereichte pragmatische Information ($A_{|prag:P}$), übersetzt wird in das wie im Index spezifizierte Sign TA, unter demselben Kontext C. C ist natürlich zunächst die Satz- oder auch Text-DRS. C kann aber auch Teil-Semantiken daraus in den Blick nehmen, vgl. Abschnitt 6.

Die zweite Regel wirkt entsprechend für Head-Complement-Strukturen. In beiden Fällen ist der entscheidende Punkt wie die pragmatische Information nach unten gereicht wird und die semantische Target-Information, $TDRS$, nach oben.

4 Basisregeln: Transfer lexikalischer Knoten

- $\boxed{\text{ich}}$

$$\text{tau}(S_{[\text{drs: } <\text{i}, \boxed{\begin{array}{c}\text{i}\\ \hline \\ \end{array}}_{<\ i,\text{sprecher}>}>]}, T_{[\text{drs: } <\text{t(i)}, \boxed{\begin{array}{c}\text{t(i)}\\ \hline \\ \end{array}}_{<\ t(i),\text{sprecher}>}>]}, C).$$

- $\boxed{\text{Sie}}$

$$\text{tau}(S_{[\text{drs: } <\text{u}, \boxed{\begin{array}{c}\text{u}\\ \hline \\ \end{array}}_{<\ u,\text{hörer}>}>]}, T_{[\text{drs: } <\text{t(u)}, \boxed{\begin{array}{c}\text{t(u)}\\ \hline \\ \end{array}}_{<\ t(u),\text{hörer}>}>]}, C).$$

- $\boxed{\text{morgen}}$

$$\text{tau}(S_{\left[\begin{array}{ll}\ldots\text{mod:} & <^{\text{M}}[\text{ drs: } <\epsilon, D_\epsilon>]> \\ \text{drs:} & <\epsilon, \boxed{\begin{array}{l}\text{t n}\\ \hline \text{next\_day(n)=t}\\ \epsilon\oslash\text{ t}\end{array}}^{\cup D_\epsilon}_{<\text{n,sprechzeit}>} >\end{array}\right]}, T_{\left[\text{drs: } <\text{t}(\epsilon), \boxed{\begin{array}{l}\text{t(t) t(n)}\\ \hline \text{next\_day(t(n))=t(t)}\\ \text{t}(\epsilon)\ \oslash\text{t(t)}\end{array}}^{\cup D_{t(\epsilon)}}_{<\text{t(n),sprechzeit}>} >\right]}, C)$$
$$\leftarrow M_{|prag:P}\ \&\ \text{tau}(M, TM_{[\text{drs: } <\text{t}(\epsilon), D_{t(\epsilon)}>]}, C).$$

Die notierten Transfer-Lexikoeinträge sollten relativ selbsterklärend sein. Bei den DRSen zu Sprecher und Hörer findet nur Variablenumbenennung statt. [5] *Morgen* als Adjunkt ist ein Ereignistyp-Modifikator. Dh. es existiert ein Head-Feature *mod*, das einen Ereignistyp, $<\epsilon, D_\epsilon>$, erwartet, wobei

[5] Es ist klar, daß bei einer echten Variablenumbenennung eine Buchhaltung geführt werden muß, die es gestattet, für den Transfer selber, aber auch im Rahmen der Dialog-Semantik, bei der Anaphorik etc. vom DRF zu seinem Bild zu gelangen und umgekehrt. Echte Umbenennung und Buchhaltung ist allerdings nur dann notwendig, wenn unter DRFs strukturierte Objekte verstanden werden, die Information über ihre sprachspezifische Realisierung wie Genus, Numerus etc. enthalten, wie momentan in Verbmobil für die Objekte im Universum einer DRS üblich.

morgen die Lokationszeit des ausgezeichneten DRFs dieses Typs, ϵ, einschränkt. Es wurde hier eine eher analytische Form des semantischen Beitrags von *morgen* gewählt - es wird der Folgetag zum Sprechzeitpunkt eingeführt, der mit ϵ überlappen muß - so daß die relativ sprachferne Notation direkt in die Target-DRS übernommen werden kann. Bis auf die Variablenumbenennung ein Fall also von lokalem shared content. Rekursiv muß allerdings noch die Target-Struktur zum modifizierten Sign errechnet werden.

- $\boxed{\text{treffen}}$

$$\mathrm{tau}(S\begin{bmatrix} \text{prag:} & \text{Abschluß} \\ \ldots\text{subcat:} & <\ [\text{drs:} < x, D_x >], ②[\text{ drs:} < y, D_y >]> \\ \text{drs:} & <\text{e}, \begin{array}{|l|}\hline \text{e} \\ \hline \text{e: } \boxed{\text{treffen(x,y)}} \\ \text{theme1(e)=x} \\ \text{theme2(e)=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{bmatrix}, T\begin{bmatrix} \text{prag: Abschluß} \\ \text{drs:} \quad <\text{t(e)}, \begin{array}{|l|l|}\hline \text{t(e)} & \text{t(x)} \\ \hline \multicolumn{2}{|l|}{\text{t(e): } \boxed{\text{see(t(x),t(y))}}} \\ \multicolumn{2}{|l|}{\text{theme1(t(e))=t(x)}} \\ \multicolumn{2}{|l|}{\text{theme2(t(e))=t(y)}} \\ \hline \end{array} \cup D_{t(y)} > \end{bmatrix}, C)$$

$$\leftarrow D_x = \begin{array}{|c|}\hline \text{x} \\ \hline \\ \hline \end{array}_{< x,\text{sprecher}>} \;\&\; D_y = \begin{array}{|c|}\hline \text{y} \\ \hline \\ \hline \end{array}_{< y,\text{hörer}>} \;\&\; \mathrm{tau}(②, T2_{\left[\text{drs: } < t(y), D_{t(y)} >\right]}, C).$$

$$\mathrm{tau}(S\begin{bmatrix} \text{prag:} & \text{P} \\ \ldots\text{subcat:} & < ①[\text{ drs:} < x, D_x >], ②[\text{ drs:} < y, D_y >]> \\ \text{drs:} & <\text{e}, \begin{array}{|l|}\hline \text{e} \\ \hline \text{e: } \boxed{\text{treffen(x,y)}} \\ \text{theme1(e)=x} \\ \text{theme2(e)=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{bmatrix}, T\begin{bmatrix} \text{prag: P} \\ \text{drs:} \quad <\text{t(e)}, \begin{array}{|l|}\hline \text{t(e)} \\ \hline \text{t(e): } \boxed{\text{meet(t(x),t(y))}} \\ \text{theme1(t(e))=t(x)} \\ \text{theme2(t(e))=t(y)} \\ \hline \end{array} \cup D_{t(y)} \cup D_{t(x)} > \end{bmatrix}, C)$$

$$\begin{array}{l} \leftarrow ①_{|prag:P} \;\&\; ②_{|prag:P} \;\& \\ \mathrm{tau}(①, T1_{\left[\text{drs: } < t(x), D_{t(x)} >\right]}, C) \;\&\; \mathrm{tau}(②, T2_{\left[\text{drs: } < t(y), D_{t(y)} >\right]}, C). \end{array}$$

Für *treffen* finden wir zwei Einträge im Transfer-Lexikon. Der erste handelt einen Spezialfall ab: Am Gesprächsende kann *treffen*, wenn in der Konstellation "Sprecher wendet sich damit an Hörer - also: *ich treffe Sie* ...", auch mit *see* übersetzt werden. Genau diese Bedingungen sind durch den pragmatischen Constraint und die Prämissen der Regel formuliert. Treffen diese Bedingungen zu, kann in der Übersetzung auch auf die Realisierung der Rolle aus dem Subjekt, hier tentativ *theme1* genannt, verzichtet werden. Dem ist in der Transfer-Regel dadurch Rechnung getragen, daß $t(x)$ durch die Abtrennung im Universum der Target-DRS als implizites Argument ausgezeichnet wird und der rekursive Aufruf von *tau* für die Subjektsposition unterdrückt wird. Die zweite Regel handelt den Normalfall ab. Innerhalb der Transfer-Routine ist natürlich die Anwendung einer spezifischen Regel der Anwendung einer allgemeineren vorzuziehen.

Bezogen auf Beispiel (1) erhalten wir den folgenden Ablauf: Die Übersetzung des Gesamt-Signs wird durch die Adjunct-Head-Perkolationsregel auf die Übersetzung des *morgen*-Knotens zurückgespielt. Die passende Regel aus dem Transfer-Lexikon verlangt dann in Rekursion die Übersetzung des *ich treffe Sie*-Knotens. Die Head-Complement-Perkolationsregel führt diesen Fall entlang der an diesem Knoten endenden semantischen Projektionsline zurück auf die Übersetzung des *treffe*-Knotens. Die hier passende Regel aus dem Transfer-Lexikon verlangt dann in letztem Neuaufruf von *tau* die Übersetzung der subkategorisierten Funktionen.[6]

Man beachte, daß Selektionsbeschränkungen lokal sehr schön abgeprüft werden können, weil das Semantische Prinzip dafür sorgt, daß Rollen am Verbknoten gefüllt sind. Ähnliches gilt für andere Funktortypen, wie Adjunkte. Es kann aber, durch das Weiterreichen des Kontexts auch andere Information abgefragt werden.

[6] Transferrekursion über die Argumente von Prädikaten findet sich auch schon in Ansätzen wie Zajac(1989) Russell et al.(1991), die den hier beschriebenen Ansatz inspiriert haben.

Der folgende Abschnitt soll an einem Beispiel aus dem Referenzdialog Ökonomie und Flexibilität der vorgeschlagenen Transfer-Prozedur deutlich machen.

5 Transfer am Beispiel von DE008

DE008:

- *Das passt echt schlecht bei mir.*

$$\begin{bmatrix} \text{prag: 'replik,neg.bew.proposition'} \\ \text{drs:} \begin{array}{|l|}\hline now_2, u_{2,pro}, s_2, i_2 \\ \hline s_2\!: \boxed{\text{echt(schlecht)(passen)}(u_2,i_2)} \\ \text{theme}(s_2)=u_2 \\ \text{exp}(s_2)=i_2 \\ now_2 \subset s_2 \\ \hline \end{array} \begin{array}{l} < now_2,\text{sprechzeit}_2 > \\ < i_2,\text{sprecher}> \end{array} \\ \ldots\text{subcat: } <> \end{bmatrix}$$

Die gegebene Repräsentation weicht von dem korrespondierenden Vorschlag in Bos/McGlashan(1994) ab. Sie ist bewußt sprachspezifischer gehalten als in dem dortigen Vorschlag – wie in der Tendenz sämtliche Repräsentationen zum Referenzdialog aus der Sicht des Transfers (Eberle(1994)) eher sprachspezifisch sind verglichen mit Bos/McGlashan(1994). Nach Copestake et al.(1994)) wird für die Entwicklung der englischen Grammatik, also zumindest auf der Targetseite, solche Sprachnähe gewünscht. Verbleibt man bei der Source-Seite bei einer relativ sprachfernen Repräsentationsform wird der Transfer, wie Tests zeigen, auch damit umgehen können, bei genügender Flexibilität was die Target-Seite anlangt.

In unserer Repräsentation von DE008 wir *echt* als Adverbmodifikator behandelt, ergibt also angewandt auf ein Adverb wieder ein Adverb. Adverben wie *schlecht* können offensichtlich nicht extensional intersektiv behandelt werden (im Sinne $e\!:\boxed{passen(x,y)} \wedge schlecht(e)$), allgemeiner sind sie deshalb Operatoren die Ereignistypen in Ereignistypen überführen. Bei *schlecht* wird dabei, extensional betrachtet der Input-Ereignistyp auf eine Teilmenge reduziert. *echt(schlecht)* sondert, bei selbem Input aus dieser Teilmenge eine kleinere aus, also: $\lambda x, y(\text{echt(schlecht)(passen)}(x, y)) \subseteq \lambda x, y(\text{schlecht(passen)}(x, y)) \subseteq \lambda x, y\ (\text{passen})(x, y))$.
Bild 4 zeigt die (wieder um Stellungsphänomene vereinfachte) Töchterstruktur zu DE008. Es ist auf die Saarbrücker *alpha*-Kondition zur Markierung anaphorischer Ausdrücke verzichtet worden. Es wurde eine eher konventionelle Notationsweise(wie ohne Hinweis schon in den Abschnitten zuvor), mit Indices an den DRFs, gewählt, die eine flachere DRS ermöglicht. Ein wesentlicher Punkt ist, daß die bei-PP hier kein Adjunkt ist, sondern von *passen* subkategorisiert wird. [7]

[7] In Verbmobil wird dies bisher anders behandelt. Wichtiger als daß die Subkategorisierung unserer Sichtweise entspricht ist aber an dieser Stelle, daß diese Variante in einfacher Weise die ökonomische Funktionsweise des Transfer-Algorithmus veranschaulichen kann.

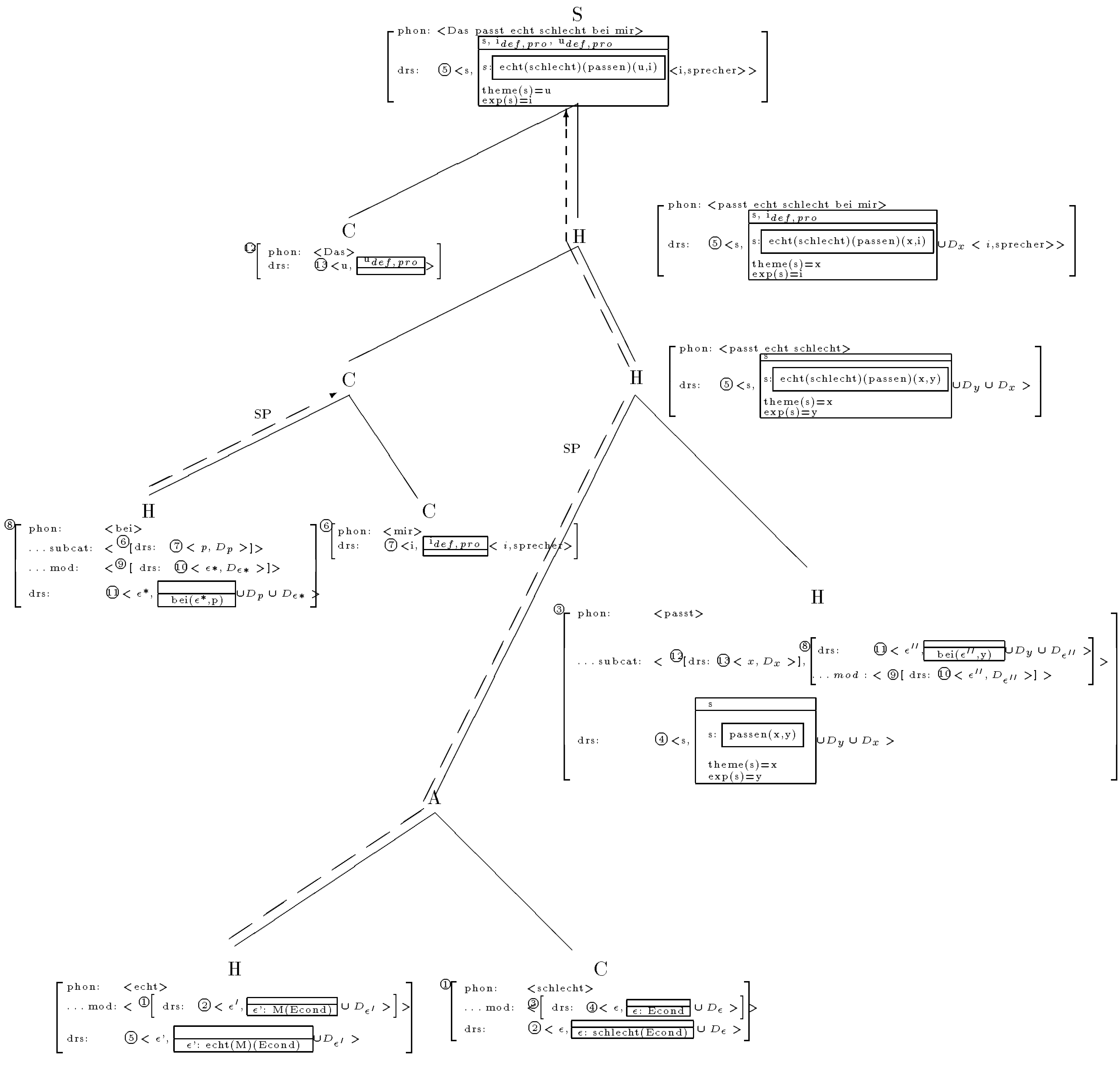

Abbildung 4: Die Töchterstruktur für DE008

5.1 Weitere Basisregeln: Transfer lexikalischer Knoten

Im folgenden notieren wir die für DE008 benötigten Basisregeln, aber schon strukturiert als prototypische Vertreter von jeweils ganzen Klassen.

5.1.1 Regeln für Pronomen (und Nomen)

- R_{pro1}

$$\mathrm{tau}(S_{\left[\ \mathrm{drs:}\ <\mathrm{i},\ \begin{array}{|l|}\hline {}^{\mathrm{i}}{def,pro} \\ \hline \\ \hline \end{array} < i,\mathrm{sprecher}>\ \right]}, T_{\left[\ \mathrm{drs:}\ <\mathrm{t(i)},\ \begin{array}{|l|}\hline {}^{\mathrm{t(i)}}{def,pro} \\ \hline \\ \hline \end{array} < t(i),\mathrm{sprecher}>\ \right]}, C).$$

- R_{pro2}

$$\mathrm{tau}(S_{\left[\ \mathrm{drs:}\ <\mathrm{x},\ \begin{array}{|l|}\hline \mathrm{x}_{pro} \\ \hline \\ \hline \end{array}\right]}, T_{\left[\ \mathrm{drs:}\ <\mathrm{t(x)},\ \begin{array}{|l|}\hline \mathrm{t(x)}_{pro} \\ \hline \\ \hline \end{array}\right]}, C)$$

$\leftarrow$*"die Beschreibung von x in S erlaubt eindeutige Pronomengenerierung in der Zielsprache ohne weitere Inferenz"* & $S_{|drs.ind.descr} = T_{|drs.ind.descr}$.

- R_{pro3}

$$\mathrm{tau}(S_{\left[\ \mathrm{drs:}\ <\mathrm{x},\ \begin{array}{|l|}\hline \mathrm{x}_{pro} \\ \hline \\ \hline \end{array}\right]}, T_{\left[\ \mathrm{drs:}\ <\mathrm{t(x)},\ \begin{array}{|l|}\hline \mathrm{t(x)}_{pro} \\ \hline \mathrm{t(x)=t(y)} \\ \hline \end{array}\right]}, C)$$

$\leftarrow(C \vdash_{DoM} \mathrm{x=y})$ & *"y nicht pronominal (dh. ist Argument einer Kondition)"*.

- R_{noun1}

$$\mathrm{tau}(S_{\left[\ \mathrm{drs:}\ <\mathrm{X},\ \begin{array}{|l|}\hline \mathrm{X} \\ \hline \mathrm{unterlage^*(X)} \\ \hline \end{array}\right]}, T_{\left[\ \mathrm{drs:}\ <\mathrm{t(X)},\ \begin{array}{|l|}\hline \mathrm{t(X)} \\ \hline \mathrm{paper^*(t(X))} \\ \hline \end{array}\right]}, C)$$

$\leftarrow(C \vdash_{DoM} \mathrm{X{:}PAPER^*})$.

- R_{noun2}

$$\mathrm{tau}(S_{\left[\ \mathrm{drs:}\ <\mathrm{X},\ \begin{array}{|l|}\hline \mathrm{X} \\ \hline \mathrm{unterlage^*(X)} \\ \hline \end{array}\right]}, T_{\left[\ \mathrm{drs:}\ <\mathrm{t(X)},\ \begin{array}{|l|}\hline \mathrm{t(X)} \\ \hline \mathrm{desk\_pad^*(t(X))} \\ \hline \end{array}\right]}, C)$$

$\leftarrow \ldots\ldots$

R_{pro1} wiederholt den schon notierten Fall von Pronomen für die erste Person. R_{pro2} beschreibt den Fall eines Pronomens der Quellsprache, das auch ohne Kenntnis des Antezedenten eindeutig übersetzt werden kann, beispielsweise *du*. In diesem Fall erbt der Target-DRF die syntaktische Beschreibung des Source-DRF. [8] Die vage "…"-Schreibweise für restringierende Konditionen ist im Realfall natürlich zu ersetzen durch exakte Formulierungen vermöge passender Zugriffsprädikate, bzw. Schnittstellenprädikate, vor allem zur Semantischen Auswertung. R_{pro3} macht Gebrauch von **der** Schnittstelle zur Semantischen Auswertung. $C \vdash_{DoM} F$ steht für *leite die Formel F aus dem Kontext C über der Wissensbasis DoM (Domain Model) ab.* R_{pro3} verlangt speziell für DRFs aus Pronomen, die nicht eindeutig übersetzt werden können, die Identifikation mit einem nicht-pronominalen Antezedenten. [9] R_{noun1} und das nicht völlig ausgearbeitete R_{noun2} reflektieren exemplarisch Fälle bei denen nach dem BACK-Konzept eines DRF gefragt wird.[10]

5.1.2 Regeln für Verben

- R_{verb1}

$$\mathrm{tau}(S\left[\begin{array}{ll} \mathrm{prag:} & \mathrm{P} \\ \ldots\mathrm{subcat:} & < ①[\ \mathrm{drs:}\ < x, D_x >], \left[\begin{array}{ll} \mathrm{drs:} & < \epsilon, \begin{array}{|l|}\hline \\ \hline \mathrm{bei}(\epsilon,\mathrm{y}) \\ \hline \end{array} \cup D_y \cup D_\epsilon > \\ \ldots\mathrm{mod:} & < [\ \mathrm{drs:}\ < \epsilon, D_\epsilon >] > \\ \mathrm{dtrs.comp:} & ②[\ \mathrm{drs:}\ < y, D_y >] \end{array}\right] > \\ \mathrm{drs:} & <\mathrm{s}, \begin{array}{|l|}\hline \mathrm{s} \\ \hline \mathrm{s:}\ \boxed{\mathrm{passen(x,y)}} \\ \mathrm{theme(s)=x} \\ \mathrm{exp(s)=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{array}\right], T\left[\begin{array}{ll} \mathrm{prag:} & \mathrm{P} \\ \mathrm{drs:} & <\mathrm{t(s)}, \begin{array}{|l|}\hline \mathrm{t(s)} \\ \hline \mathrm{t(s):}\ \boxed{\mathrm{suit(t(x),t(y))}} \\ \mathrm{theme(t(s))=t(x)} \\ \mathrm{exp(t(s))=t(y)} \\ \hline \end{array} \cup D_{t}(y) \cup D_{t}(x) > \end{array}\right], C)$$

$$\leftarrow ①_{|prag:P}\ \&\ ②_{|prag:P}\ \&\ \mathrm{tau}(①, T1_{[drs:<t(x),D_{t(x)}>]}, C)\ \&\ \mathrm{tau}(②, T2_{[drs:<t(y),D_{t(y)}>]}, C.)$$

[8] An dieser Stelle ist natürlich wieder darüber nachzudenken, was der Status von DRFs ist, bzw. wo Information über ihre Realisierung im Satz notiert wird.

[9] Man kann, da Pronominalisierung auch eine Frage der Sprachplanung ist, solche Anfragen auch ganz der Generierung überantworten, oder umgekehrt ganz dem Transfer, auch in Fällen, wo die Ziel-Pronominalisierung eindeutig wäre.

[10] Die Alternativen für Pronomen und Common Nouns machen deutlich, wie der Transfer das Konzept der variablen Analysetiefe ausnützt. Vertiefte Analysen werden nur bei Bedarf angefordert, wobei die Semantische Auswertung gesteuert und beschränkt wird durch die Formulierung der Anfrage als spezifischer kontextueller Constraint (vgl. Eberle et al.(1992)).

In den Repräsentationen für *Unterlagen* stehen großgeschriebene Variablen für Summenreferenten, und mit * versehene Sorten S und einstellige Prädikate P für den Summenabschluß über den Extensionen von S bzw. P.

- R_{verb2}

$$\mathrm{tau}(S\begin{bmatrix} \text{prag:} & \text{P} \\ \text{...subcat:} & <\text{①}[\ \text{drs:}\ <x, D_x>],\text{②}[\ \text{drs:}\ <y, D_y>]> \\ \text{drs:} & <\text{s}, \begin{array}{|l|}\hline \text{s} \\ \hline \text{s:}\ \boxed{\text{passen(x,y)}} \\ \text{theme(s)=x} \\ \text{exp(s)=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{bmatrix}, T\begin{bmatrix} \text{prag:} & \text{P} \\ \text{drs:} & <\text{t(s)}, \begin{array}{|l|}\hline \text{t(s)} \\ \hline \text{t(s):}\ \boxed{\text{suit(t(x),t(y))}} \\ \text{theme(t(s))=t(x)} \\ \text{exp(t(s))=t(y)} \\ \hline \end{array} \cup D_t(y) \cup D_t(x) > \end{bmatrix}, C)$$

$$\leftarrow ①_{|prag:P}\ \&\ ②_{|prag:P}\ \&\ \mathrm{tau}(①, T1_{[drs:<t(x),D_{t(x)}>]}, C)\ \&\ \mathrm{tau}(②, T2_{[drs:<t(y),D_{t(y)}>]}, C.)$$

R_{verb1} behandelt den Fall mit subkategorisierter PP. In diesem Fall ist es unnötig, eine Übersetzung für den PP-Knoten als Ganzes anzubieten. Es genügt völlig, die von der Präposition subkategorisierte Phrase zu übersetzen. Im Gegensatz zu einem reinen bottom-up-Verfahren läßt das vorgeschlagene kopfbezogene Transferverfahren in diesem und anderen einschlägigen Fällen bestimmte Knoten bei der Übersetzung aus, dann wenn die entsprechenden Übersetzungen gar nicht gebraucht werden. Das ist ein wesentliches Moment der Ökonomie. R_{verb2} behandelt den simplen Fall von zweistellig *passen*. Es ist vielleicht an der Stelle darauf hinzuweisen, daß das Zeichen für die DRS-Vereinigung, $\cup$, bei der Implementierung der *compose*-Operation entspricht, so daß der Source-Konstruktionsalgorithmus nach unserem Ansatz auch bei der Erstellung der Target-Semantik benutzt wird.[11]

5.1.3 Regeln für Adverben

- R_{adv1}

$$\mathrm{tau}(S\begin{bmatrix} \text{prag:} & \text{P} \\ \text{...mod:} & <①\left[\ \text{drs:}\ <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: Econd} \\ \hline \end{array} \cup D_\epsilon >\right]> \\ \text{drs:} & <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: schlecht(Econd)} \\ \hline \end{array} \cup D_\epsilon > \end{bmatrix}, T\begin{bmatrix} \text{prag:} & \text{P} \\ \text{drs:} & <t, \begin{array}{|l|}\hline \text{t} \\ \hline \neg\text{: }(\begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: good(TEcond)} \\ t(\epsilon) \oslash \text{t} \\ \hline \end{array} \cup D_{t(\epsilon)}) \\ \hline \end{array}> \end{bmatrix}, C)$$

$$\leftarrow \text{Econd} = \text{passen(X,Y)}\ \&\ ①_{|prag:P}\ \&\ \mathrm{tau}(①, T1_{\left[\ \text{drs:}\ <t(\epsilon), \begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: TEcond} \\ \hline \end{array} \cup D_{t(\epsilon)} >\right]}, C).$$

- R_{adv2}

$$\mathrm{tau}(S\begin{bmatrix} \text{prag:} & \text{P} \\ \text{...mod:} & <①\left[\ \text{drs:}\ <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: Econd} \\ \hline \end{array} \cup D_\epsilon >\right]> \\ \text{drs:} & <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: schlecht(Econd)} \\ \hline \end{array} \cup D_\epsilon > \end{bmatrix}, T\begin{bmatrix} \text{prag:} & \text{P} \\ \text{drs:} & <t(\epsilon), \begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: bad(TEcond)} \\ \hline \end{array} \cup D_{t(\epsilon)} > \end{bmatrix}, C)$$

$$\leftarrow ①_{|prag:P}\ \&\ \mathrm{tau}(①, T1_{\left[\ \text{drs:}\ <t(\epsilon), \begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: TEcond} \\ \hline \end{array} \cup D_{t(\epsilon)} >\right]}, C).$$

Statt *Econd = passen(X,Y)* wird man in R_{adv1} im Realfall eine geeignete allgemeinere Bedingung einsetzen, um den Anwendungsbereich der Regel zu maximieren.

Die konditionslos gegebene Regel R_{adv2} dient als "catch all", zur Berücksichtigung aller adverbialen Verwendungen von *schlecht*, die durch R_{adv1} nicht erfasst werden. Die Anordnung entspricht der generellen Strategie zur Anwendungsreihenfolge der Regeln.

5.1.4 Regeln für Adverbmodifikatoren

- R_{advm1}

$$\mathrm{tau}(S\begin{bmatrix} \text{phon:} & <\text{echt}> \\ \text{prag:} & \text{P} \\ \text{...mod:} & <\begin{bmatrix} \text{drs:} & <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: schlecht(passen(x,y))} \\ \text{theme}(\epsilon)\text{=x} \\ \text{exp}(\epsilon)\text{=y} \\ \hline \end{array} \cup D_y \cup D_x > \\ \text{...mod:} & \left[\ \text{...subcat:}\ <①[\ \text{drs:}\ <x, D_x>], \begin{bmatrix} \text{drs:} & <\epsilon, \begin{array}{|l|}\hline \\ \hline \text{bei}(\epsilon,\text{y}) \\ \hline \end{array} \cup D_y \cup D_\epsilon > \\ \text{...mod:} & <[\text{drs:}\ <\epsilon, D_\epsilon>]> \\ \text{dtrs.comp:} & ②[\ \text{drs:}\ <y, D_y>] \end{bmatrix}>\right] \end{bmatrix}> \\ \text{drs:} & <\epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: echt(schlecht)(passen(x,y))} \\ \text{theme}(\epsilon)\text{=x} \\ \text{exp}(\epsilon)\text{=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{bmatrix},$$

[11] Der Abschnitt 7 zur Implementierung macht in der Verwendung semantischer Macros allerdings deutlich, daß die Target-Semantik keineswegs notwendig das Format der Source-Semantik haben muß. Wir haben in der Einleitung schon auf diesen Punkt hingewiesen. Es ist denkbar, daß bei entsprechendem Design der Target-Grammatik, die Interpretation dieser Macros anders sein kann als auf der Source-Seite. Wesentlich ist, zu sehen, daß auch bei verändertem Target-Format, die Formulierung der Transferregeln stabil bleibt. Darin liegt ein großer Vorteil des vorgeschlagenen Ansatzes.

$$T\left[\begin{array}{ll} \text{prag:} & \text{P\&umgangssprache} \\ \text{drs:} & < t, \begin{array}{|l|}\hline t \\ \hline \neg_{at\_all}:(\begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: suit(t(x),t(y))} \\ \text{theme(t(}\epsilon\text{))=x} \\ \text{exp(t(}\epsilon\text{))=y} \\ \text{t } \oslash t(\epsilon) \\ \hline \end{array} \cup D_{t(y)} \cup D_{t(x)}) \\ \hline \end{array} > \end{array}\right], C)$$

$$\leftarrow \textcircled{1}_{|prag:P} \;\&\; \textcircled{2}_{|prag:P} \;\&\; \text{tau}(\textcircled{1}, T1_{[drs:<t(x),D_{t(x)}>]}, C) \;\&\; \text{tau}(\textcircled{2}, T2_{[drs:<t(y),D_{t(y)}>]}, C)$$

- R_{advm2}

$$\text{tau}(S\left[\begin{array}{ll} \text{phon:} & <\text{echt}> \\ \text{prag:} & \text{P} \\ \ldots\text{mod:} & < \left[\begin{array}{ll} \text{drs:} & < \epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: schlecht(passen(x,y))} \\ \text{theme(}\epsilon\text{)=x} \\ \text{exp(}\epsilon\text{)=y} \\ \hline \end{array} \cup D_y \cup D_x > \\ \ldots\text{mod:} & \left[\begin{array}{l} \ldots\text{subcat:} < \textcircled{1}[\text{ drs: } < x, D_x >], \\ \qquad\qquad \textcircled{2}[\text{ drs: } < y, D_y >]> \end{array}\right] \end{array}\right] > \\ \text{drs:} & < \epsilon, \begin{array}{|l|}\hline \epsilon \\ \hline \epsilon\text{: echt(schlecht)(passen(x,y))} \\ \text{theme(}\epsilon\text{)=x} \\ \text{exp(}\epsilon\text{)=y} \\ \hline \end{array} \cup D_y \cup D_x > \end{array}\right], T\left[\begin{array}{ll} \text{prag:} & \text{P\&umgangssprache} \\ \text{drs:} & < t, \begin{array}{|l|}\hline t \\ \hline \neg_{at\_all}:(\begin{array}{|l|}\hline t(\epsilon) \\ \hline t(\epsilon)\text{: suit(t(x),t(y))} \\ \text{theme(t(}\epsilon\text{))=x} \\ \text{exp(t(}\epsilon\text{))=y} \\ \text{t } \oslash t(\epsilon) \\ \hline \end{array} \cup D_{t(y)} \cup D_{t(x)}) \\ \hline \end{array} > \end{array}\right], C)$$

$$\leftarrow \textcircled{1}_{|prag:P} \;\&\; \textcircled{2}_{|prag:P} \;\&\; \text{tau}(\textcircled{1}, T1_{[drs:<t(x),D_{t(x)}>]}, C) \;\&\; \text{tau}(\textcircled{2}, T2_{[drs:<t(y),D_{t(y)}>]}, C)$$

- R_{advm3}

$$\text{tau}(S\left[\begin{array}{ll} \text{phon:} & <\text{echt}> \\ \text{prag:} & \text{P} \\ \ldots\text{mod:} & < \textcircled{1}\left[\text{ drs: } < \epsilon, \begin{array}{|l|}\hline \\ \hline \epsilon\text{: M(Econd)} \\ \hline \end{array} \cup D_\epsilon >\right] > \\ \text{drs:} & < \epsilon, \begin{array}{|l|}\hline \\ \hline \epsilon\text{: echt(M)(Econd)} \\ \hline \end{array} \cup D_\epsilon > \end{array}\right], T\left[\begin{array}{ll} \text{prag:} & \text{P\&umgangssprache} \\ \text{drs:} & < t(\epsilon), \begin{array}{|l|}\hline \\ \hline t(\epsilon)\text{: really(TM)(TEcond)} \\ \hline \end{array} \cup D_{t(\epsilon)} > \end{array}\right], C)$$

$$\leftarrow \textcircled{1}_{|prag:P} \;\&\; \text{tau}(\textcircled{1}, T1_{\left[\text{ drs:} < t(\epsilon), \begin{array}{|l|}\hline \\ \hline t(\epsilon)\text{: TM(TEcond)} \\ \hline \end{array} \cup D_{t(\epsilon)} > \right]}, C).$$

R_{advm1} und R_{advm2} spezifizieren den Sonderfall, bei dem *echt* auf *schlecht(passen)* trifft. Es wird hier, anstatt rekursiv tau für das Adjunkt *schlecht* aufzurufen und damit die Übersetzungsaufgabe zu dekomponieren, eine Übersetzung für den ganzen Ausdruck *echt schlecht passen* angeboten. Es bleibt dann noch, die Rollen aus dem Subkategorisierungsrahmen zu *passen* zu übersetzen. R_{advm1} und R_{advm2} thematisieren die von den Lexikoneinträgen zu *passen* bekannten Fälle. Es ist allerdings möglich, daß zum Zeitpunkt der Adverb- und Adverbmodifikatoranwendung der ursprüngliche Subkategorisierungsrahmen schon teilweise oder ganz abgearbeitet worden ist. Diese Fälle müssen, bei einer genauen Ausarbeitung natürlich berücksichtigt werden (vgl. zu einer kompakten Berücksichtigung solcher Fälle die Funktion *taus* im Abschnitt zur Implementierung). R_{advm3} spiegelt den unspezifischen Fall wider.

5.2 Die Target-DRS

Durch Anwendung von R_{advm1} erhalten wir aus DE008 die folgende Target-DRS:

<u>DE008</u>:
- *That doesn't suit me at all.*

⇐*Das passt echt schlecht bei mir.*

$$\left[\begin{array}{ll} \multicolumn{2}{l}{\text{prag: 'replik,neg.bew.proposition,umgangssprache'}} \\ \text{drs:} & \begin{array}{|l|}\hline now_2, u_{2,pro} \\ \hline \neg_{at\_all}: \begin{array}{|l|}\hline s_2, i_{2,pro,def} \\ \hline s_2\text{:} \boxed{\text{suit}(u_2, i_2)} \\ \text{theme}(s_2)=u_2 \\ \text{exp}(s_2)=i_2 \\ now_2 \subseteq s_2 \\ \hline \end{array} \\ \hline \end{array} \begin{array}{l} < now_2,\text{sprechzeit}_2 > \\ < i_2,\text{sprecher}> \end{array} \end{array}\right]$$

Die Anwendung der unspezifischen Regel R_{advm3} ergibt dagegen das folgende [12]:

- *That really doesn't suit me.*

$\Leftarrow$*Das passt echt schlecht bei mir.*

$$\begin{bmatrix} \text{prag: 'replik,neg.bew.proposition,umgangssprache'} \\ \text{drs:} \begin{array}{|l|}\hline now_2, u_{2,pro} \\ \hline \text{really}(\neg: \begin{array}{|l|}\hline s_2, i_{2,pro,def} \\ \hline s_2: \boxed{\text{suit}(u_2, i_2)} \\ \text{theme}(s_2)=u_2 \\ \text{exp}(s_2)=i_2 \\ now_2 \subset s_2 \\ \hline \end{array}) \\ \hline \end{array} \begin{array}{l} < now_2, \text{sprechzeit}_2 > \\ < i_2, \text{sprecher} > \end{array} \end{bmatrix}$$

6 Besonderheiten

- R_{verb3}

$$\text{tau}(S \begin{bmatrix} \text{prag:} & \text{P} \\ \text{...subcat:} & < ①[\text{ drs: } < x, D_x >], ②[\text{ drs: } < y, D_y >]> \\ \text{drs:} & <\text{s}, \begin{array}{|l|}\hline \text{s, p} \\ \hline \text{s: } \boxed{\text{glauben(x,p)}} \\ \text{theme1(s)=x} \\ \text{theme2(s)=p} \\ \text{p: } D_y \\ \hline \end{array} \cup D_x > \end{bmatrix}, T \begin{bmatrix} \text{prag: P} \\ \text{drs:} \quad <\text{t(s), t(p)}, \begin{array}{|l|}\hline \text{t(s)} \\ \hline \text{t(s): } \boxed{\text{believe(t(x),t(p))}} \\ \text{theme1(t(s))=t(x)} \\ \text{theme2(t(s))=t(p)} \\ \text{t(p): } D_{t(y)} \\ \hline \end{array} \cup D_{t(x)} > \end{bmatrix}, C)$$

$$\leftarrow ①_{|prag:P} \ \&\ \text{tau}(①, T1_{[drs:<t(x), D_{t(x)}>]}, C) \ \&\ \text{tau}(②, T2_{[drs:<t(y), D_{t(y)}>]}, [D_y, C].)$$

Normalerweise wird die Pragmatikinformation zur Übersetzung von Argumenten oder modifizierten Strukturen nach unten weitergegeben. Es gibt aber bestimmte einbettende DRS-Konditionen, die für die eingebettete DRS eine andere pragmatische Bewertung wahrscheinlich machen als für die Gesamt-DRS. Ein Beispiel hierfür sind bestimmte Einstellungskontexte, wie das in R_{verb3} thematisierte *Glauben*. Die Regel verlangt hier zwar das Weiterreichen des Pragmatikwerts zur Analyse des Glaubenden, aber nicht das Weiterreichen zum Glaubensinhalt. Die Glaubenseinstellung ist auch ein Beispiel für Kontextveränderung im dritten Argument der tau-Regel. Soll etwas für die Ebene des Glaubensinhalts bewiesen werden, muß, für diesen Zweck, diese Ebene als real betrachtet werden. Das leistet die Heraushebung des Glaubensinhalts D_y aus der eingebetteten Position zum neuen Kontext $[D_y, C]$. Im Zusammenhang mit einer DRS-Beweistheorie (vgl. Reyle(1994)) muß die Kontextveränderung an Fallbeispielen allerdings weiter ausdifferenziert werden. Motivation hierfür ist, der Semantischen Auswertung Hilfe zu leisten aus einer Position heraus, bei der klar ist für welche DRS-Ebene etwas abgeleitet werden soll. Das ist eine Information, die der Semantischen Auswertung selber zunächst nicht bekannt ist.

7 Implementierung

Zur Realisierung des sign-basierten Transfers wird im folgenden die Übersetzungsrelation `tau` zwischen Signs aus der Quell- und Zielsprache definiert. Aufgrund der funktionalen Schreibweise von Prädikaten in STUF wird die 2-stellige Relation `tau(sign,sign)` [13] als `tau(sign) => sign` notiert,

[12] Die bisher eingeführten Regeln haben hauptsächlich illustrativen Charakter. Um mit R_{advm3} tatsächlich die folgende DRS zu erreichen müssen weitere Details eingearbeitet werden. Es fehlen auch Tempus- und Aspektregeln, deren Wirkung in der DRS aber schon berücksichtigt ist.

[13] In der momentanen Implementierung wurde eine Vereinfachung bzgl. des globalen Kontextes vorgenommen. Das bedeutet, daß kein zusätzlicher Parameter in der Relation verwendet wird und der globale Kontext - sollte dies nötig

vgl. die entsprechende Typdeklaration in STUF:

```
@tau(sign) => sign.
```

Bei der Verbmobil-Implementierung ist die Ausgabe der Analyse eine Liste von Signs. Die Argumente eines Prädikats werden als Liste notiert, genauso wie das Modificandum eines semantischen Operators. Deshalb gibt es eine weitere Relation `taus`, die eine Liste von Signs mit Hilfe der `tau`-Relation in eine Liste von Signs für die Zielsprache übersetzt. Diese Relation wird in substantieller Weise bei der Übersetzung der Prädikatsargumente verwendet.

```
@taus(list) => list.
taus([]) => [].
taus([#F|#R]) => [tau(#F)|taus(#R)].
```

Die Gesamtmenge der Transferregeln `tau`, die für die Übersetzung von Quell- in Ziel-Signs benötigt werden, lassen sich aufteilen in eine Menge von rekursiven Regeln, die beschreiben in welcher Form die Dekomposition der Quell-DRS erfolgen soll, und in eine Menge von Basisregeln `tau_lex`, die die Korrespondenzen für den lexikalischen Transfer beschreiben.

Die Formulierung von rekursiven Transferregeln als Rekursion allein auf der Quell-DRS erweist sich als problematisch, da die Konditionenmenge einer DRS nicht unmittelbar die Ableitungsgeschichte für die Konstruktion der DRS wiederspiegelt. In dem hier vorgeschlagenen Ansatz wird deshalb die hierarchische Struktur des phrasalen Skeletts im Quell-Sign entlang semantischer Projektionslinien benutzt, um rekursive Transferregeln formulieren zu können. Die Definition semantischer Köpfe ergibt sich unmittelbar aus der Formulierung des Semantischen Prinzips der zugrundeliegenden HPSG-Grammatik. Die Rekursion entlang der semantischen Kopfprojektion wird beendet, sobald ein Blatt mit lexikalischer Information (also vom Typ `word_s`) erreicht wird. In diesem Falle wird die Relation `tau_lex` (s.u.) aufgerufen:

```
tau(word_s & #Sign) => tau_lex(#Sign).
```

Handelt es sich bei dem Quell-Sign um ein phrasales Sign (also vom Typ `phrase_s`), dann sind abhängig von der lokalen Konstituentenstruktur unterschiedliche Fälle für die Bestimmung der semantischen Kopfprojektion zu berücksichtigen:

1. Grundsätzlich läßt sich unterscheiden zwischen Kopf-Argument-Strukturen (vom Typ `head_comps_struc`) und nicht Kopf-Argument-Strukturen.

   (a) Adjunkte verhalten sich als semantische Funktoren, die Rekursion von `tau` folgt daher der Adjunkt-Tochter (Wert des Merkmals `adj_dtr`):

   ```
   tau(phrase_s &
       dtrs: (head_adj_struc &
              adj_dtr: #SemHead)) => tau(#SemHead).
   ```

   (b) Gleiches gilt für ein topikalisiertes Adjunkt (Wert des Merkmals `top_dtr`), das ebenfalls als semantischer Kopf fungiert.

   ```
   tau(phrase_s &
       dtrs: (head_topic_struc &
              top_dtr: #SemHead)) => tau(#SemHead).
   ```

werden - durch geeignetet Zugriffsprädikate angefragt werden kann. Der lokale Kontext steht durch das Quell-Sign natürlich zur Verfügung. Grundsätzlich ist noch unklar, welcher Kontext für Inferenzen benötigt bzw. ausgewählt werden muß. Was die Weitergabe der Pragmatik-Information anbelangt, wird bei der Implementierung davon ausgegangen, daß diese Information vom globalen Kontext subsummiert wird - Kontext also DRS **und** Pragmatik ist - und die im letzten Abschnitt angesprochene Frage von Barrieren für Pragmatik und DRS-Kontext einheitlich über Zugriffsprädikate zu regeln ist, so, allerdings, daß im Ergebnis die im letzten Abschnitt beschriebene Funktionalität entsteht.

(c) Für alle anderen nicht Kopf-Argument Strukturen (Head-Filler- und Head-Conjuncts-Struktur) gilt, daß der semantische Kopf gleich dem syntaktischen Kopf (Wert des Merkmals `h_dtr`) ist.

```
tau(phrase_s &
    dtrs: (( head_filler_struc
           ; head_conjs_struc) &
           h_dtr: #SemHead)) => tau(#SemHead).
```

2. Bei den Kopf-Argument-Strukturen sind ebenfalls mehrere Fälle zu unterscheiden, da die momentane syntaktische Analyse in Verbmobil für Verberst- und Verbzweitstellung eine Verbbewegung annimmt.[14]

(a) Handelt es sich bei dem syntaktischen Kopf um ein Verb in Erst- oder Zweitstellung, das aus seiner Basisposition wegbewegt wurde, dann ergibt sich die semantische Kopfprojektion über das Komplement, um sicherzustellen, daß etwaige Adjunkte im Mittelfeld vor dem finiten Verb konsultiert werden.

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   phon:cons &
                   sent_type(v1;v2)) &
           c_dtrs: [#SemHead])) => tau(#SemHead)
```

(b) In allen anderen Fällen, ist der semantische Kopf gleich dem syntaktischen Kopf, wie es gemäß dem Semantikprinzip der HPSG zu erwarten ist.

i. Verbletztstellung oder nicht-verbale lexikalische Projektion:

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   #SemHead &
                   phon:cons &
                   sent_type(v3)))) => tau(#SemHead).
```

ii. Phrasale Projektion:

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (phrase_s  &
                   #SemHead))) => tau(#SemHead).
```

iii. Spur:

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   #SemHead &
                   phon:[]))) => tau(#SemHead).
```

Durch die soeben definierte Relation `tau` kann offensichtlich die phrasale Struktur eines Signs durchwandert werden, um den semantischen Kopf zu finden. Auf diesen wird die zweistellige Relation `tau_lex` appliziert.

[14] Der Korrektheit wegen sei erwähnt, daß für die folgenden Definitionen der `tau`-Relation syntaktische Kategorien zur feineren Fallunterscheidung benötigt werden. Diese sind z.B. nötig, um eine Instantiierung nicht-verbaler Knoten durch die Relation `sent_type` zu verhindern. Leider stehen Kategorien momentan (noch) nicht zur Verfügung.

`tau_lex` bestimmt zum einen die Übersetzung lexikalischer Signs der Quellsprache in (nicht notwendig lexikalische) Signs der Zielsprache, zum anderen die weitere Rekursion anhand der Prädikat-Argument-Struktur des Bild-Signs bzw., allgemeiner, anhand von dessen semantischer Funktor-Agument(e)-Struktur. Im einfachen Fall reduziert sich die Formulierung lexikalischer Transferregeln zu der Aufgabe, eindeutige Prädikatsnamen (siehe `sem_lex` weiter unten) der Quell- und Zielsprache miteinander in Beziehung zu setzen und geeignete Bedingungen für die Anwendbarkeit dieser Relation in Abhängigkeit vom lokalen bzw. globalen Kontext zu formulieren. Auf diesen einfachen Fall konzentrieren wir uns im folgenden bei der Illustration von `tau_lex`-Regeln. Die Implementierung von Fällen, die eine komplexere Spezifikation der Targetsemantik notwendig machen, wie beispielsweise die phrasalen Übersetzungen von spezifischen Vorkommen von *echt* in Abschnitt 5.1.4, wird gesondert in einem Papier zu typischen Übersetzungsschwierigkeiten vorgestellt werden.

Die Relation `tau_lex` extrahiert den Prädikatsnamen aus dem Quell-Sign mit Hilfe des neu einzuführenden Schnittstellenprädikats `pred_name` und übergibt diesen an das erste Argument der vierstelligen Relation `tau_lex`, in der die eigentlichen lexikalischen Transferregeln definiert sind. Bei erfolgreicher Anwendung dieser Relation wird automatisch der Zielprädikatsname im zweiten Argument instantiiert und an die semantische Zugriffsrelation `sem_lex` für die Zielsprache übergeben.

```
@tau_lex(sign) => sign.
tau_lex(pred_name(#SourcePred)&#Sign) =>
        tau_lex(#SourcePred,#TargetPred,#Sign) &
        sem_lex(#TargetPred).
```

Im folgenden sind eine Reihe von einfachen Beispielen zur Illustration prototypischer Einträge aufgeführt, die z.B. für die Übersetzung von Verbmobil-Dialogphrasen wie *"das paßt bei mir"* in *"that suits me"* oder *"wir treffen uns"* in *"we will meet"* bzw. in *"see you"* verwendet werden können.[15]

```
@tau_lex(atom,atom,sign) => sign.

/* verbs */
tau_lex(treffen,see,#SourceSign & pred_args([#SArg &
              anchor(discourse_role: sprecher)|#Args]) &
              prag:dact:abschluss) => pred_args(taus(#Args)).
tau_lex(treffen,meet1,#SourceSign & pred_args([#SArg,arg: reciprocal]) &
              prag:dact:vorschlag) => pred_args([#SArg]).
tau_lex(treffen,meet2,#SourceSign &
              prag:dact:vorschlag) => tau_args_args(#SourceSign).
tau_lex(passen,suit,#SourceSign) => tau_args_args(#SourceSign).
tau_lex(bei,suit,#SourceSign &
              syn:mod([pred_name(passen)])) => tau_prep_verb(#SourceSign).
/* adverbs */
tau_lex(dann, then, #SourceSign) => tau_mod_mod(#SourceSign).
/* prepositions */
tau_lex(bei,at,#SourceSign) => tau_prep_prep(#SourceSign).
tau_lex(in,at,#SourceSign &
              pred_concept(in, !in_time_c)) => tau_prep_prep(#SourceSign).
tau_lex(in,in,#SourceSign &
              pred_concept(in, !in_location_c)) => tau_prep_prep(#SourceSign).
/* pronouns */
```

[15] Die Definitionen der `tau_lex`-Klausel für *passen* und der zugehörigen `tau_prep_verb`-Relation sind durch die etwas unkonventionelle Analyse von *bei* als Adjunkt-Kopf in Verbindung mit *passen* innerhalb der momentanen Verbmobil-Syntax nötig. Unter der Annahme, daß *"Das paßt bei mir."* und *"Das paßt mir."* die gleiche semantische Struktur zeigen, sollte es sich nicht um ein Adjunkt, sondern vielmehr um ein Argument von *passen* handeln. Die entsprechende Transferregel wäre dann um ein Vielfaches verständlicher.

```
tau_lex(pron,pron,alfa_expr(#Alfa) &
              anchor(#Anchor)) => alfa_expr(#Alfa) & anchor(#Anchor).
tau_lex(das,that,eps_expr(#Eps)) => eps_expr(#Eps).
/* nouns */
tau_lex(treffen1,meeting,#SourceSign) => ##.
```

In obigen Definitionen von `tau_lex` wurden neue semantische Zugriffsprädikate wie `pred_concept` oder `alfa_expr` nötig, auf die hier jedoch nicht eingegangen wird, da sie eher die Syntax-Semantik-Schnittstelle betreffen als den Transfer an sich. Es ist klar, daß die Definitionen nur illustrativ, nicht ausschöpfend spezifiziert sind. Zum Beispiel wird in der Übersetzung von *meet* in *see* am Dialogende mit dem Sprecher als Wert der Subjektsfunktion gezeigt, wie in der Übersetzung das Subjekt weggelassen werden kann. Ansonsten werden zwei Formen von *meet* als Übersetzungen von *treffen* vorgesehen, das einstellige *meet1* für *treffen* mit Reflexivpronomen und das zweistellige *meet2* für *treffen* mit unterschiedlichen Referenten für Subjekt und Objekt. Natürlich sind auch andere Modellierungen denkbar. Bei der Übersetzung von *in* in *at* wird man außer der temporalen Lesart der Präposition bei der realen Regel die Punktualität des internen Arguments oder Ähnliches verlangen.

Was noch gezeigt werden soll sind Relationen, die die Abbildung von Argumenten eines Prädikats oder Operators des Quell-Signs auf die eines Ziel-Signs vornehmen. Diese sind dann für den weiteren Transfer der Argumente verantwortlich, was man an den Aufrufen der Relation `taus` erkennen kann.

```
@tau_args_args(sign) => sign.
tau_args_args(pred_args(#Args)) => pred_args(taus(#Args)).
@tau_mod_mod(sign) => sign.
tau_mod_mod(syn:mod(#M)) => syn:mod(taus(#M)).
@tau_prep_prep(sign) => sign.
tau_prep_prep(#SourceSign) => tau_mod_mod(#SourceSign) &
                              tau_args_args(#SourceSign).
@tau_prep_verb(sign) => sign.
tau_prep_verb(pred_args([#PArg]) &
              syn:mod([pred_args([#SArg|#Args])])) =>
    pred_args(taus([#SArg,#PArg|#Args])).
```

Um eine sinnvolle Indizierung der lexikalischen Transferlexikoneinträge vornehmen zu können, gehen wir davon aus, daß die Syntax-Semantik-Schnittstelle eindeutige Prädikatsnamen für die lexikalischen Signs im semantischen Lexikon zur Verfügung stellt. Die Prädikatsnamen dienen dabei als einfache Schnittstelle zwischen Analyse, Transfer und Generierung, da sie als Index Zugriff auf den semantischen Beitrag des jeweiligen Lexemes erlauben. Dabei sollte sichergestellt werden, daß unterschiedliche semantische Repräsentationen unterschiedliche Prädikatsnamen erhalten (beispielsweise das einstellige *meet* den Namen *meet1*, das zweistellige den Namen *meet2*), während semantisch mehrdeutige Lexeme durchaus den gleichen Prädikatsnamen als Index erhalten können (beispielsweise die in-Präposition den Namen *in*), solange sie die gleiche semantische Repräsentation zugewiesen bekommen. Für das semantische Lexikon hätte dies den weiteren Vorteil, daß keine Einträge für Vollformen geschrieben werden müßten, sondern ausschließlich Einträge für die eindeutigen Prädikatsnamen. Gleiches gilt für die semantischen Einträge im englischen Lexikon, auf die uniform über die Relation `sem_lex` zugegriffen werden kann, wenn eine semantische Repräsentation für ein Prädikt erzeugt werden soll. Für die Instantiierung bestimmter Werte in dieser semantischen Struktur sind dabei u.U. noch semantische Zugriffsrelationen nötig, die jedoch sowieso für den Transfer zur Verfügung gestellt werden müßten. Exemplarisch könnten diese `sem_lex`-Klauseln wir folgt aussehen, wobei wir uns an den bisher bereits existierenden Makros des Mini-Demo-Systems orientiert haben:

```
@sem_lex(atom)    => sign.
sem_lex(see)     => transitive(agent,theme,see,ereignis).
```

```
sem_lex(meet2)   => transitive(agent,theme,meet2,ereignis).
sem_lex(suit)    => transitive(agent,theme,suit,ereignis).
sem_lex(at)      => prep(at,perspective).
sem_lex(pron)    => pronoun(person).
sem_lex(that)    => pronoun(epsilon).
```

Makros expandieren zu spezifischen (partiellen) DRSen, das mit `agent,theme,meet2,ereignis` instantiierte `transitive(Role1,Role2,Pred,Sort)` expandiert zu einem Ausdruck mit der Bedeutung

$$\lambda < y, D_y >, \lambda < x, D_x > \boxed{\begin{array}{l} \text{e: ereignis} \\ \hline \text{e: } \boxed{\text{meet2(x,y)}} \\ \text{agent(e)=x} \\ \text{theme(e)=y} \end{array}} \oplus D_x \oplus D_y,$$

wobei in Attribut-Wert-Notation der relevante Strukturausschnitt der folgende ist:

```
(word_s &
 sem:(sem_t &
      lambda: ...
      ind:(#A & named_var & sort:ereignis) &
      drs:(drs_t &
           dom:[(marker & ref: #A)] &
           conds:[(basic_cond & pred:meet2 & inst: #A &
                   args:[(arg_roles & arg: #B & role:agent),
                         (arg_roles & arg: #C & role:theme)])]) &
      quants:[]))
```

Da uns weder semantische noch syntaktische Strukturen der englischen Ziel-Grammatik vorliegen, bilden wir momentan (zu Demonstationszwecken) auf die bereits innerhalb des Mini-Demonstrators implementierten Strukturen ab. Als Beispiel für die Transfer-Ausgabe des Satzes *"Wir treffen uns dann."* erhält man damit das folgende Ziel-Sign (die weitestgehend unterspezifizierte syntaktische Struktur wurde ausgeblendet):

```
(phrase_s &
 sem:(sem_t &
      lambda:[] &
      ind:(#A & named_var & sort:ereignis) &
      drs:(drs_t &
           dom:[(marker & ref: #A)] &
           conds:[(alfa_expr &
                   alfa_arg:(#B & named_var & number:pl & sort:person) &
                   alfa_type:std &
                   alfa_restr:(drs_t &
                               dom:[(marker & ref: #B)] &
                               conds:[])),
                  (basic_cond & pred:meet & inst: #A &
                   args:[(arg_roles & arg: #B & role:agent)]),
                  (basic_cond & pred:then & inst: #A &
                   args:[(arg_roles &
                          arg:(#D & named_var & sort:zeit) &
                          role:int_arg)]),
                  (alfa_expr &
                   alfa_arg: #D &
                   alfa_type:std &
                   alfa_restr:(drs_t &
                               dom:[(marker & ref: #D)] &
                               conds:[]))]) &
      quants:[] &
      anchors:[(anchor & param: #B & discourse_role:sprecher_hoerer)]))
```

8 Resümee

Die vorgeschlagene Transferkomponente übersetzt HPSG-Strukturen mit quellsprachspezifischer semantischer Repräsentation in unterspezifizierte HPSG-Strukturen mit zielsprachspezifischer Repräsentation. Zur Behandlung von Übersetzungsproblemen auf der Ebene der Bedeutungsstruktur wird vorausgesetzt, daß für die verschiedenen Sprachen des MÜ-Systems (Deutsch, Englisch und zukünftig auch Japanisch) Axiomensysteme die sprachspezifischen Prädikate an das sprachunabhängig konzipierte Domänenmodell, vor allem an die Konzepthierarchie, anbinden. Anstatt nun aber, zur Auflösung lexikalischer translation mismatches, generelle Navigationsverfahren über die, durch die Prädikatsanbindung existierende multilinguale Klassenhierarchie zu formulieren (vgl. Barnett et al.(1994), Kameyama et al.(1991)), werden spezifische lexikalische Regeln formuliert, die, bei Bedarf, nach dem Konzept der variablen Analysetiefe (vgl. etwa Kay et al.(1994)) von der semantischen Auswertung durch die Formulierung kontextueller Constraints (vgl. Eberle et al.(1992)) gezielt eine tiefere Analyse anfordern (der typische Fall etwa R_{noun1} und R_{noun2} in Abschnitt 5.1.1). Durch die vorgeschlagene Architektur ist es möglich, wie in Abschnitt 5.1.4 skizziert, größere Übersetzungseinheiten in einer einzigen Regel zu erfassen. Wir werden demnächst in einer systematischen Aufarbeitung einschlägiger Übersetzungsschwierigkeiten zeigen, wie im vorliegenden Ansatz u.a. damit Probleme im Zusammenhang mit idiomatischen Ausdrücken, Funktionsverbgefügen, Head Switching, Kategoriewechsel u.ä. (vgl. etwa Erbach/Krenn(1993), Kuhn/Heid(1994)) einer adäquaten Lösung zugeführt werden.

Literatur

Barnett, J./Mani, I./Rich, E. (1994): Reversible Machine Translation: What to do when the languages do not match up. In: Strzalkowski, T. (ed.), *Reversible Grammar in Natural Language Processing*, chapter 13, pp. 321–364. Kluwer Academic Publishers, Dordrecht, 1994

Bos, J./E.Mastenbroek/S.McGlashan/S.Millies/M.Pinkal (1994): The Verbmobil Semantic Formalism. VM report 6, Universität des Saarlandes, 1994

Bos, J./McGlashan, S. (1994): Extensions to the Verbmobil Semantic Formalism and Analysis of the Referenzdialog. VM memo 43,94, Universität des Saarlandes, 1994

C.J.Rupp/Rohrer, C./Eberle, K./Schiller, A. (1994): Lexicalised Transfer for the VERBMOBIL Demonstrator. (ms.) IMS, Universität Stuttgart, 1994

Copestake, A./Flickinger, D./Malouf, R./Riehemann, S./Sag, I. (1994): Transfer and Minimal Recursion Semantics. ms., CSLI, Stanford, 1994

Eberle, K./Kasper, W./Rohrer, C. (1992): Contextual Constraints for MT. In: *4th International Conference on Theoretical and Methodological Issues in Machine Translation*, Montreal, 1992

Eberle, K. (1991): Ereignisse: Ihre Logik und Ontologie aus textsemantischer Sicht. IWBS Report 192, IBM Deutschland, WT LILOG, Stuttgart, 1991. Wiederabdruck der Dissertation

Eberle, K. (1994): Sprachnahe DRSen des Referenzdialogs: Semantischer Transferinput und -output aus der Sicht der Transferkomponente. (ms.), IMS, Universität Stuttgart, 1994

Erbach, G./Krenn, B. (1993): Idioms and Support-Verb Constructions in HPSG. Claus-Report 28, Universität des Saarlandes, 1993

Kameyama, M./Ochitani, R./Peters, S. (1991): Resolving Translation Mismatches with Information Flow. In: *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, 1991

Kay, M./Gawron, J. M./Norwig, P. (1994): *VERBMOBIL: A Translation System for Face-to-Face Dialog.* CSLI, Stanford, 1994

Kuhn, J./Heid, U. (1994): Treating Structural Differences in an HPSG-Based Approach to Interlingual Machine Translation. In: *Tagungsband der Jahrestagung der DGfS*, Münster, 1994

Pollard, C./Sag, I. (1994): *Head-Driven Phrase Structure Grammar.* Chicago University Press, Chicago, 1994

Reyle, U. (1994): Underspecified Discourse Representation Structures and their Logic. In: Gabbay, D./Kempson, R. (eds.), *Proof Theory and Natural Language* Special Issue of the IGPL. 1994

Russell, G./Ballim, A./Estival, D./Warwick-Armstrong, S. (1991): A Languge for the Statement of Binary Relations over Feature Structures. In: *Fifth Conference of the European Chapter of the Association for Computational Linguistics*, pp. 287–292, Berlin, 1991

Whitelock, P. (1992): Shake-and-Bake-Translation. In: *Proceedings of the 14th International Conference on Computational Linguistics (Coling-92*, Nantes, France, 1992

Zajac, R. (1989): A Transfer Model Using a Typed Feature Structure Rewriting System with Inheritance. In: *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 1–6, Vancouver, 1989