Robuste Semantik in Verbmobil

Schriftliche Ausarbeitung zum Referat im Proseminar
„Maschinelle Übersetzungssysteme"
Dr. Paul Schmidt
SS 2000

Garance PARIS

3. Dezember 2000

Sprach- und Translationswissenschaft /
Maschinelle Übersetzung

Fachrichtung 5.6 Angewandte Sprachwissenschaft
sowie Übersetzen und Dolmetschen

Universität des Saarlandes

Inhaltsverzeichnis

Einleitung

Das Projekt Verbmobil bestand aus zwei Phasen, von 1993 bis 1996 und von 1997 bis 2000. Nach der ersten Phase wurde im Verbmobil-Forschungsprototyp bemerkt, daß in vielen Fällen keine überspannende Analyse für die ganze Äußerung möglich war: unzusammenhängende erkannte Fragmente wurden einfach aneinandergereiht und ohne weiteres der Transferkomponente überreicht.

Dieses Problem, unter dem die Übersetzungsqualität leidet, hat mehreren Ursachen, die mit der Verarbeitung von Spontansprache zu tun haben und in Abschnitt 1 beschrieben werden.

Daher die Besonderheit der semantischen Verarbeitung in Verbmobil: Sie muß rrobustßein, um zusammengehörige, durch Fehler getrennte Fragmente zu größeren Einheiten kombinieren zu können, bevor diese dann dem Transfer weitergegeben werden. Der Begriff der „Robustheit" wird in Abschnitt 2 zunächst theoretisch erläutert, dann wird an Hand von Beispielen gezeigt, wie diese Robustheit bei der zweiten Phase des Projektes realisiert wird.

Diese Arbeit referiert zum größten Teil den Artikel von K. Worm (1997). Wenn keine Literaturangaben im weiteren explizit genannt werden, stammt alles aus dieser Quelle.

1 Probleme der Verarbeitung von Spontansprache

Bei der Verarbeitung von Spontansprache erhält das semantische Modul als Eingabe Sprachobjekte, die aus verschiedenen Gründen keine grammatische, vollständige Sätze darstellen, und die daher zu keiner überspannenden semantischen Analyse führen können. Die Problemquellen für die semantische Verarbeitung kann man in vier Klassen aufteilen:

- Selbstkorrekturen & Abbrüche der Sprecher,
- Erkennerfehler,
- Segmentierungsprobleme, und
- Unzureichende Abdeckung der Grammatik.

1.1 Selbstkorrekturen & Abbrüche der Sprecher

1.1.1 Definitionen: Selbstkorrektur und Abbruch

Da eine Implementation beabsichtigt wird, benötigt man genaue Definitionen. Die hier benötigten Definitionen lauten wie folgt:

Selbstkorrektur: "Der Sprecher korrigiert innerhalb eines Turns einen Teil der bisherigen Äußerung [...] An einer Stelle unterbricht der Sprecher die Äußerung und setzt sie nach einer u. U. durch Interjektionen gefüllten Pause fort und äußert etwas, was einen Teil der Äußerung vor der Unterbrechung ersetzt."

Abbruch: "Der Sprecher bricht eine Äußerung ab [...] und beginnt eine neue [...] Die abgebrochene Äußerung ist meistens [...] syntaktisch unvollständig [...] wenn sie einen vollständigen Satz bildet, muß anhand der Prosodie erkennbar sein, daß die Äußerung unvollständig ist."

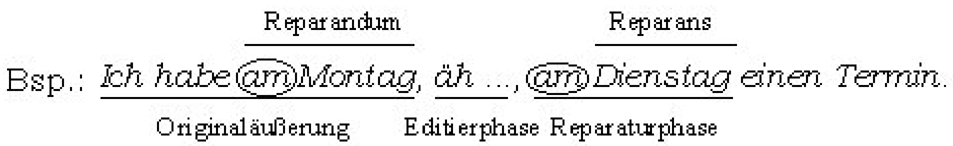

Abbildung 1.1: Beispiel einer Selbstkorrektur

1.1.2 Ein Modell für Selbstkorrektur

Es werden verschiedene Modelle der Selbstkorrektur vorgeschlagen, hier reicht aber zu sagen, daß eine Selbstkorrektur aus drei Teilen besteht:

- einer Originaläußerung, die das Reparandum enthält,
- einer Editierphase, die aus einer Pause besteht, eventuell gefüllt mit Interjektionen und Korrekturmarker (*'äh', 'ich meine', 'also', ...*),
- einer Reparaturphase, bestehend aus Reparans und manchmal mit Teilwiederholung des Reparandums.

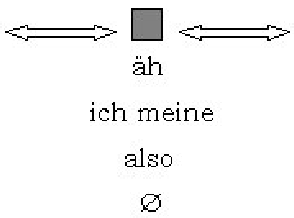

Abbildung 1.2: Ein Modell für Selbstkorrektur

Außerdem werden Selbstkorrekturen in bestimmten Fällen auch durch eine unregelmäßige prosodische Grenze markiert. Dies hängt allerdings davon ab, wie stark korrigiert wird: Die Markierung ist häufiger und wird mehr betont, wenn der Sprecher eine falsche Angabe verbessert, als wenn er eine andere Formulierung wählt. Bei Abbrüchen dient Valenzinformation zur Erkennung, unterstützt von prosodischer Information.

1.2 Erkennerfehler

Der Spracherkenner bildet aus dem akustischen Sprachsignal eine Kette von Wörtern. Trotz Fortschritten in der Spracherkennung gibt es noch viele Fehler: Wörter werden nicht erkannt oder verschluckt, usw. Daher entspricht das Ergebnis nur in seltenen Fällen dem, was tatsächlich geäußert wurde, und es ist auch häufig nicht syntaktisch wohlgeformt und daher nicht analysierbar.

Aus diesem Grunde werden mehrere Hypothesen als Eingabe für die weitere Verarbeitung ausgewählt, um die Wahrscheinlichkeit zu erhöhen, daß wenigstens eine richtig ist. Es dürfen aber auch nicht zuviele ausgewählt werden: Gute Ergebnisse bei der Spracherkennung müssen gegen den späteren Aufwand bei der syntaktischen Verarbeitung abgewogen werden, denn mit der Anzahl der Hypothesen steigt gleichzeitig auch der Suchaufwand für den Parser.

Diese Hypothesen werden nach den akustischen Bewertungen geordnet, die vom Spracherkenner geliefert wurden. Dies wird als Worthypothesengraph dargestellt (s. Abbildung 1.3).

Doch ist die der Äußerung entsprechenden Kette nicht immer unter der Menge der Hypothesen und es ist auch nicht sicher, daß überhaupt eine analysierbare Kette in der Menge der Hypothesen ist.

1.3 Segmentierungsprobleme

In der gesprochenen Sprache gibt es keine Interpunktion, daher muß zur Segmentierung die Prosodie dienen. Die Segmentierung erfolgt in Verbmobil auf verschiedenen Ebenen: Zunächst ermittelt die prosodische Analyse mögliche Phrasen- und Satzgrenzen, dann werden diese Grenzen angenommen oder verworfen, abhängig davon, ob eine Analyse für den Segment gefunden werden konnte oder nicht.

Wenn keine überspannende Analyse möglich war, wird das Segment weiter aufgeteilt und es wird eine Folge kleineren Segmenten erstellt.

Dies führt zu Problemen, wenn Fragmente, die nicht zusammengehören, in einem Segment zusammengefaßt werden, oder umgekehrt, Teile getrennt werden.

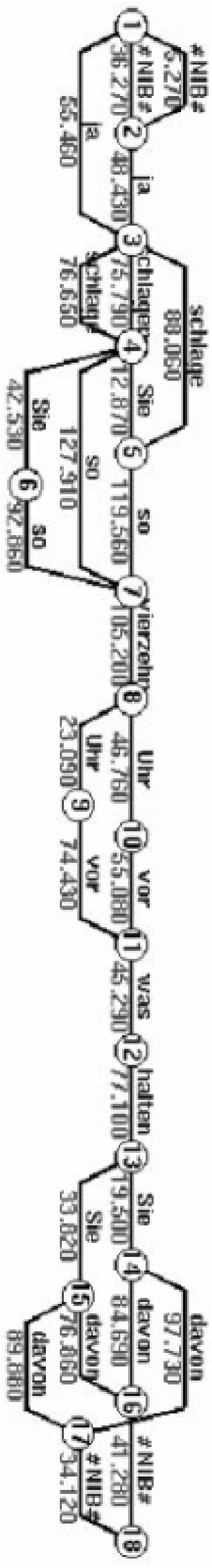

Abbildung 1.3: Worthypothesengraph für die Äußerung:
„Ja ich schlage so 14 Uhr vor. Was halten Sie davon?"
(NIB = Nicht interpretierbarer Bereich)

1.4 Unzureichende Abdeckung der Grammatik

Die Grammatiken, die für natürlichsprachliche Systeme eingesetzt werden, können nicht alle grammatischen Konstruktion vorgesehen haben. Dies ist bei gesprochener Sprache in besonderer Maße wahr, da hier außerdem Konstruktionen, die in geschriebener Sprache unzulässig sind, und Anakoluthe auftauchen.

Die Abdeckung der Grammatik ist nicht nur wegen spontansprachlicher Probleme zu klein, sondern es gibt auch absichtliche Beschränkungen: Es ist bei natürlichsprachlichen Systemen sinnvoll, seltene Konstruktionen herauszulassen, um den Suchraum für den Parser zu verkleinern, weil das Durchprobieren aller möglichen Kombinationen sehr zeit- und speicherintensiv werden kann.

2 Robuste semantische Verarbeitung

2.1 Systemarchitektur

Die Übersetzung in Verbmobil erfolgt mit einem Transferansatz. Die semantische Verarbeitung setzt also nach der syntaktischen Analyse an und operiert auf der Ausgabe der verschiedenen linguistischen Ebenen.

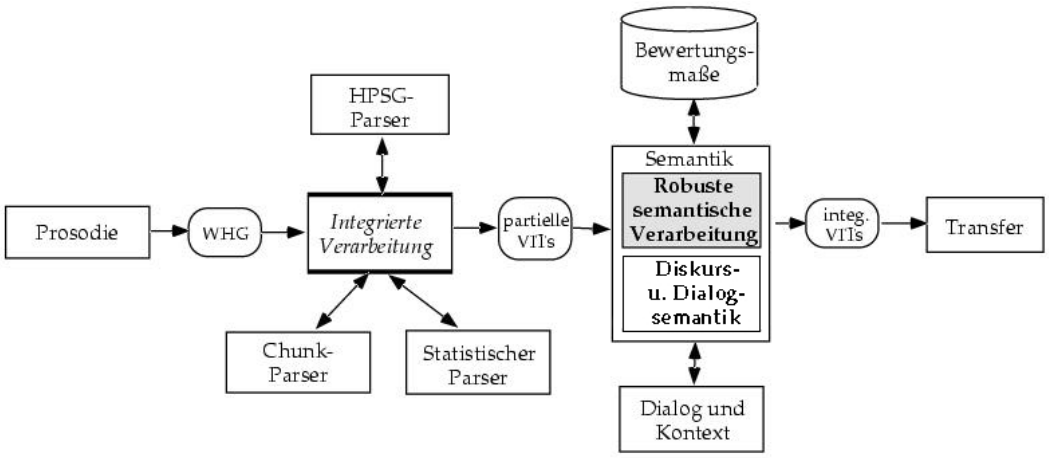

Abbildung 2.1: Systemarchitektur

2.2 Aufgabe der robusten Semantik

Die Semantik besteht aus zwei Modulen, 'Diskurs- und Dialogsemantik' und 'Robuste semantische Verarbeitung'. Dies liegt daran, daß eine einfache diskurs- und dialogsemantische Verarbeitung wegen der im ersten Abschnitt erwähnten Probleme nicht reicht, die großenteils an der Verarbeitung von Spontansprache liegen.

Um die Qualität der Analyse und dadurch der Übersetzung zu steigern, müssen solche Probleme überwunden werden: In der künstlichen Intelligenz wird von „Robustheit" oder auch von „Fehlertoleranz" gesprochen.

Insbesondere müssen robuste semantische Verfahren verwendet werden, die:

- bruchstückhafte Analysen zu größeren Einheiten kombinieren, und
- Selbstkorrekturen und Abbrüche herausfiltrieren.

Dabei sind keine Wunder zu erwarten: Das Ergebnis hängt immer davon ab, was die robuste Semantik als Eingabe bekommt. In Fällen, in denen wesentliche bedeutungstragende Teile auf Grund von Fehlern bei der vorherigen Analyse nicht vorhanden sind, kann nicht sehr viel verbessert werden.

2.3 Robuste semantische Verfahren

2.3.1 Grundlegende Konzepte

Verbmobil Interface Terms

Eingabe- und Ausgabeformat der robusten Semantik sind jeweils Verbmobil Interface Terms (VITs). Dies ist eine im Rahmen von Verbmobil entwickelte Repräsentation linguistischer Information aus verschiedenen Ebenen der Verarbeitung: prosodische, morphosyntaktische und semantische.

Jeder VIT entspricht einem Segment eines Pfades auf einem Worthypothesengraph. Dieses standardisierte Format wird an den verschiedenen Schnittstellen zwischen den Modulen verwendet, um die Kommunikation zwischen den Modulen zu vereinfachen und sie effizienter zu machen (Dorna: 2000).

Das Pool-Konzept

Statt Modul-zu-Modul-Kommunikation wie in der ersten Phase von Verbmobil, wurde in der zweiten Phase ein Pool-Konzeptïmplementiert: Jedes Modul legt seine Ergebnisse in eine Datenstruktur, Pool"genannt, ab, auf die jedes andere Modul zugreifen kann. Es handelt sich also um eine offene Architektur, in der das produzierende Modul nicht wissen muß, von wem seine Ergebnisdaten später verwendet werden, und das empfangende Modul nichts mit dem Produzent explizit vereinbaren muß.

Ein Anytime-Verfahren

Zusammen mit den in VITs segmentierte Repräsentation von Äußerungen erlaubt das Pool-Konzept eine inkrementelle Verarbeitung (Dorna: 2000).

Außerdem wird aus Effizienzgründen eine Chart verwendet, ähnlich wie bei dem Parsen, um Zwischenergebnisse zu speichern, falls man sie später wieder braucht. Dieses Abspeichern von Zwischenergebnissen ermöglicht es, eine Folge partieller Analysen als Ergebnis zu bekommen, selbst wenn keine überspannende Analyse gefunden wird.

In der künstlichen Intelligenz spricht man von Änytime-Verfahrenfür solche Verfahren, die jederzeit ein Ergebnis bereit haben, wenn sie unterbrochen werden, in der aber die Qualität des Ergebnisses mit der Zeit steigt (Russell: 1995, S. 844, Wahlster: 2000, S. 305).

2.3.2 Regel- und Anwendungsbeispiele

Die Verarbeitung basiert auf heuristische Regeln, die bottom-up angewendet werden. In den Regeln wird bestimmt, welche VITs zusammengefügt oder gelöscht werden müssen.

Unter den mehreren, in der Chart vorhandenen, möglichen Ergebnissen wird dann von einer Bewertungsfunktion eins gewählt mit als Kriterien:

- die akustische Bewertung des auf dem Worthypothesengraph verfolgten Pfads,
- die Länge der Abdeckung,
- die Saturiertheit des Prädikats,
- die Verletzung von Sortenrestriktionen, usw.
- der Bonus oder Malus bei bestimmten Regeln (s. Beispiel 2).

Es folgen jetzt einige Beispiele um zu zeigen, wie die im ersten Abschnitt beschriebenen Probleme durch das Einführen solcher Regeln auf der semantischen Ebene behoben werden können.

Beispiel 1: Selbstkorrektur eines Modifikators

Selbstkorrekturen kann man nicht mit zusätzlichen Regeln in der Grammatik behandeln, weil der Suchraum dadurch zu sehr vergrößert wäre. Sie müssen also auf höherer Ebene behandelt werden, nämlich in der Semantik.

Da bei Terminverhandlungen Zeitangaben häufig vorkommen, kommen sie auch entsprechend häufig bei Korrekturen vor. Zum Beispiel (die Beispiele sind transliterierte Daten aus dem „Verbmobil-Corpus spontansprachlicher Dialoge"):

`sagen wir am ersten Juni äh am ersten Juli`

Angenommen, es wird eine irreguläre prosodische Grenze am Ende des ersten Segments erkannt und der Parser liefert die folgenden Segmente:

`[sagen wir ersten Juni] [ersten Juli]`

Hier kommen die Definitionen und das Modell aus Abschnitt 1.1 zum Einsatz. Es wird eine Regel eingeführt, die besagt, daß in solchen Fällen, wenn die beiden Modifikatoren temporal sind und beide Datumsangaben sind, das Reparandum durch das Reparans ersetzt werden kann.

Beispiel 2:
Eine Selbstkorrektur, die zu einem ungebundenen Argument führt

Transliteration:

`das ist ... passt mir sehr gut`

Angenommen, daß bei der Spracherkennung nichts schief läuft, wird beim Parsen die zweite Verbphrase zusammengebaut. In der ersten („ist") fehlt aber das Prädikativum, daher kann die unvollständige Verbphrase von dem HPSG-Parser nicht weiter mit ihrem Subjekt kombiniert werden. Die erkannten Fragmente sind:

`[das] [ist] [passt mir sehr gut]`

Die vorige Regel läßt sich nur auf Modifikatoren anwenden, daher wird das erste Verb nicht herausgenommen. Es gibt aber eine semantische Regel, die sagt: „Wenn eine VP und eine passende NP im Nominativ gefunden werden, dann füge sie zu einem Satz zusammen". Diese Regel kann zweimal angewendet werden:
1) einmal auf „das" und „ist", wobei die Bewertung des Ergebnisses berücksichtigen muß, daß

- adjazente Fragmente kombiniert werden (positiv),
- aber daß ein Argument ungebunden bleibt (negativ), und
- daß sich hinter ïstëine irreguläre prosodische Grenze befindet (negativ),

2) und einmal auf „das" und „passt mir sehr gut", wobei hier

- eine Kante übersprungen wird (negativ),
- diese aber ist kurz (positiv)
- und endet mit einer irregulären prosodischen Grenze (positiv),
- das Ergebnis saturiert ist.

Schließlich wird von der Bewertungsfunktion die zweite Variante gewählt.

Beispiel 3: Eine fehlende Rolle wird mit Dialogakten gefüllt

Anhand von Dialogakten können in manchen Fällen ungebundene Argumentrollen durch Sprecher oder Hörer gefüllt werden.

Transliteration:

`ja ich schlage so vierzehn uhr vor was halten sie davon`

Erkannte Wortkette:

`ja schlagen so vierzehn uhr vor was halten sie davon`

Das Subjekt wurde nicht erkannt und fehlt. Angenommen, „schlagen ... vor" wird vom Parser richtig erkannt als das Verb „vorschlagen", eine Dialogaktregel könnte lauten: „Wenn ein Verb des Vorschlagens erkannt wird, aber das Subjekt fehlt, dann füge 'ich' (den Sprecher) als Subjekt ein".

Beispiel 4: Erkennerfehler

Erkannte Fragmente:

`[mir waere es am liebsten] [den naechsten zwei wochen]`

Die Präposition „in" wurde zwischen den Fragmenten nicht erkannt. Bei der Terminvereinbarung kommen Zeitadverbiale häufig vor, daher wird die folgende Regel eingeführt: „Wenn das erste Fragment ein Satz ist und das zweite eine nicht-angebundene temporale Nominalphrase ist, dann mache die Nominalphrase zu einem Zeitadverbial."

Beispiel 5:
Segmentierungsprobleme und unzureichende Abdeckung der Grammatik

Dieses Beispiel kann für beide Problemklassen aus den Abschnitten 1.3 und 1.4 dienen. Es handelt sich um zusammengehörende Fragmente, die auf einen niedrigeren Verarbeitungsebene getrennt wurden, und von der robusten Semantik wieder zusammengefaßt werden müssen.

Erkannte Fragmente:

`[thursday] [I am out]`

Es gibt zwei möglichen Gründen, warum die Äußerung in zwei Fragmente getrennt worden sein könnte: Entweder liegt dies an der Prosodie nach dem Modifikator, die von einer Komma in geschriebener Sprache wiedergegeben wird (vgl. 1.3), oder an der Abwesenheit einer Topikalisierungsregel in der Grammatik (vgl. 1.4).

Ein alleinstehender Wochentag wird in dieser Domäne wahrscheinlich als Modifikator geparst. Daher kann eine Regel eingeführt werden, die diesen Modifikator mit der Proposition kombiniert.

Zusammenfassung

Um die Qualität der Analyse und dadurch der Übersetzung zu erhöhen, werden die im ersten Abschnitt beschriebenen spontansprachlichen Phänomene in einem Modul "Robuste Semantik" behandelt.

Innerhalb dieses Moduls basiert die Verarbeitung auf heuristische Regeln, die bottom-up angewendet werden:

- Regeln, die Selbstkorrekturen mit verschiedenen Ansätzen herausfiltern (s. Beispiele 1 und 2),
- Regeln, die fehlende - d.h. nicht erkannte - Teile an Hand Textanalytischer Regeln (s. Beispiel 3) und Diskursstruktur (s. Beispiel 4: Dialogakten) ergänzen,
- Regeln, die aus sonstigen Gründen auf einen niedrigeren Verarbeitungsebene fehlerweise getrennte Fragmente in der robusten Semantik wieder zusammengefassen (Beispiel 5).

An Hand dieser und weiterer Regeln werden fragmentarische, im VIT-Format vorhandenen Ergebnisse verschiedener Parsing Ansätzen in einer Chart-Datenstruktur zu größeren Einheiten kombiniert, bevor sie dem Transfer überreicht werden, der sie in zielsprachliche VITs übersetzt.

Quellenverzeichnis

DORNA, Michael
2000 A Library Package for the Verbmobil Interface Term
Verbmobil-Report 238
(updated version of Verbmobil-Report 104)
Mai 2000
IMS, Universität Stuttgart
http://verbmobil.dfki.de/cgi-bin/verbmobil/htbin/
decode.cgi/share/VM-depot/FTP-SERVER/vm-reports/report-238-00.ps.gz

RUSSELL, Stuart J. & Norvig, Peter
1995 Artificial Intelligence, A Modern Approach
Upper Saddle River, New Jersey: Prentice Hall
Prentice Hall Series in Artificial Intelligence

WAHLSTER, W.
2000 Einführung in die Methoden der KI
Skript zur Vorlesung
Sommersemester 2000
Universität des Saarlandes, Fachbereich Informatik

WORM, Karsten L.
1997 Robuste semantische Verarbeitung
Verbmobil-Report 200
Dezember 1997
Universität des Saarlandes, Computerlinguistik
http://verbmobil.dfki.de/cgi-bin/verbmobil/htbin/
decode.cgi/share/VM-depot/FTP-SERVER/vm-reports/report-200-97.ps.gz