

# Natürlichsprachliche Beschreibung von Straßenverkehrsszenen durch Bildfolgenauswertung

Zur Erlangung des akademischen Grades eines  
Doktors der Naturwissenschaften  
der Fakultät für Informatik der Universität Karlsruhe (TH)

*genehmigte*

**Dissertation**

von

**Ralf Gerber**

aus Wehr/Baden

Erster Gutachter: Prof. Dr. H.-H. Nagel

Zweitgutachter: Prof. Dr. A. Schmitt

Prof. Dr. H. Kamp, Universität Stuttgart



# Danksagung

Mein Dank gilt in erster Linie meinem Doktorvater, Herrn Professor Dr. Hans-Hellmut Nagel. Irgendwann im Herbst 1995 hatte Herr Professor Nagel wohl beschlossen, daß er aus mir einen ‘Doktor’ macht, und von diesem guten Vorsatz konnte ihn nichts und niemand abbringen, am allerwenigsten ich selbst. Ich danke Herrn Professor Nagel für seine Konsequenz, das in mich gesetzte Vertrauen und die wohldosierte, gute und sehr lehrreiche Betreuung.

Ein weiterer Dank geht an meinen Korreferenten Herrn Professor Dr. Alfred Schmitt für die Durchsicht meiner Arbeit und die daraus resultierenden Verbesserungsvorschläge.

Aus dem gleichen Grund danke ich auch meinem zweiten Korreferenten, Herrn Professor Dr. Hans Kamp vom Institut für Maschinelle Sprachverarbeitung an der Universität Stuttgart. Ihm und seinen Mitarbeitern Herrn Michael Schiehlen und Herrn Peter Krause danke ich außerdem für ihr Interesse an meiner Arbeit und die gewinnbringenden fachlichen Diskussionen.

Dank gebührt auch all meinen ehemaligen und aktuellen Kolleginnen und Kollegen am Institut für Algorithmen und Kognitive Systeme für ihre Unterstützung und das kontinuierlich sehr gute Arbeitsklima. Im einzelnen und in alphabetischer Reihenfolge sind dies die Damen und Herren Ulla Dietrich, Angela Gehrke, Dr. Michael Haag, Marianne Hardt, Gabriele Kliebe, Holger Leuck, Markus Middendorf, Klaus Mück und Thomas Müller. Besonderer Dank gilt den Herren Haag und Leuck für ihre Unterstützung in der Endphase meiner Dissertation.

Schließlich gebührt ein dickes Dankeschön meinen Eltern für 30 Jahre Hoffen und Bangen und Sich-Sorgen-Machen.



## Kurzfassung

Fortgeschrittene Bildfolgenauswertungssysteme berechnen sehr reichhaltige Mengen an geometrischen und begrifflichen Auswertungsdaten, so daß es für den Menschen immer schwieriger wird, diese Datenfülle zu überschauen. Eine mögliche Form der Darstellung, bei der Auswertungsergebnisse effizient durch den Menschen inspiziert werden können, besteht in der Formulierung eines natürlichsprachlichen Textes.

Die vorliegende Arbeit beschäftigt sich mit der automatischen Generierung natürlichsprachlicher Beschreibungen zeitveränderlicher Szenen durch Bildfolgenauswertung. Dies wird am Beispiel des Diskursbereichs ‘Straßenverkehr’ demonstriert. Ausgangspunkt bilden digitisierte Grauwertbildfolgen von Kreuzungs- und Tankstellenszenen, die mit einer stationären Standardvideokamera aufgezeichnet wurden. Die im Rahmen der Arbeit entwickelte Systemkomponente basiert ausschließlich auf realen Auswertungsdaten, die durch das modellgestützte Bildfolgenauswertungssystem XTRACK automatisch extrahiert worden sind.

Abläufe in Szenen lassen sich stets auf mehrere Arten beschreiben, die sich inhaltlich (z.B. in der Schwerpunktsetzung und der Beurteilung) voneinander unterscheiden können. Jede Beschreibung ist somit ‘perspektivisch’ und verweist damit auf einen individuellen Standpunkt des Beschreibenden. Basierend auf linguistischen Forschungsarbeiten wird im Rahmen meiner Arbeit ein Ansatz zur generischen Modellierung des abstrakten Begriffs der *Perspektivität* eingeführt.

Die automatisch gewonnenen geometrischen Beschreibungen und das rechnerintern repräsentierte Modellwissen werden integriert und genutzt, um basierend auf einem vorhandenen, unscharf metrischen temporalen Logiksystem generisch formulierte Situationsschemata automatisch auszuprägen. So wird für jeden Handlungsträger dessen aktuelle Situation bestimmt. Im nächsten Schritt werden begriffliche, perspektivische Beschreibungen der generierten Situationen extrahiert, wobei diese Beschreibungen sich auf den Standpunkt beziehen, der sich aus Benutzer- oder Spezifikationsvorgaben ableiten läßt.

Anhand eines bestehenden logikbasierten Ansatzes der Diskursrepräsentationstheorie wird die begriffliche, perspektivische Beschreibung in eine rechnerinterne, semantische Repräsentation überführt. Aus dieser Darstellung werden natürlichsprachliche Beschreibungen in englischer Sprache erzeugt.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>12</b>
1.1	Aufgabenstellung . . . . .	12
1.2	Sprachverarbeitung in der Informatik . . . . .	15
1.3	Literaturüberblick . . . . .	16
<b>2</b>	<b>Natürlichsprachliche Beschreibungen</b>	<b>22</b>
2.1	Inhalt und Stil . . . . .	22
2.2	Erzählstil . . . . .	26
2.3	Perspektivität . . . . .	30
2.3.1	Phänomenologische Herangehensweise . . . . .	31
2.3.2	Logische Struktur der Perspektive . . . . .	32
2.3.3	Perspektive und Erzählstil . . . . .	33
2.3.3.1	Bericht . . . . .	33
2.3.3.2	Szenische Beschreibung aus Beobachtersicht (SBB) . . . . .	34
2.3.3.3	Szenische Beschreibung aus Protagonistensicht (SBP) . . . . .	35
2.4	Erzählercharakteristiken . . . . .	35
2.5	Natürlichsprachliche Anfragen . . . . .	36
2.6	Systemaufbau . . . . .	39
2.7	Geometrische Systemschicht . . . . .	41
<b>3</b>	<b>Begriffliche Abstraktion</b>	<b>44</b>
3.1	Literaturüberblick . . . . .	44
3.2	Unschärf metrisch temporale Logik . . . . .	45
3.3	Situationsanalyse . . . . .	48
3.4	Perspektivische Situationsbeschreibung . . . . .	51

3.4.1	Bestimmung des Erzählerstandpunkts . . . . .	51
3.4.2	Situationsbestimmung . . . . .	51
3.4.3	Perspektivische Beschreibungen von Situationsfolgen . . . . .	55
3.4.4	Automatisierung . . . . .	57
3.5	Erweiterte Schlußfolgerungen . . . . .	60
3.5.1	Ausnutzung von zusätzlichem Kontextwissen . . . . .	60
3.5.2	Quantorisierte Geschehen . . . . .	62
3.5.2.1	Unscharfe Quantoren . . . . .	63
3.5.2.2	Zahlwörter . . . . .	68
<b>4</b>	<b>Natürlichsprachliche Schicht</b>	<b>69</b>
4.1	Literaturüberblick . . . . .	69
4.2	Linguistische Grundlagen . . . . .	72
4.3	Diskursrepräsentationstheorie (DRT) . . . . .	76
4.3.1	Definition von DRSen . . . . .	76
4.3.2	Syntaxanalyse . . . . .	79
4.3.3	Ereignisse und Zustände . . . . .	79
4.3.4	Zeit . . . . .	84
4.4	Einbettung der DRT in Bildfolgenauswertung . . . . .	85
4.4.1	Faktensyntax . . . . .	86
4.4.2	Eigennamen . . . . .	90
4.4.3	Definite Beschreibungen . . . . .	93
4.4.4	Personalpronomen . . . . .	96
4.4.5	Prädikation . . . . .	97
4.4.6	Plural . . . . .	98
4.4.7	Indefinite Beschreibungen . . . . .	104
4.4.8	Zeit . . . . .	107
4.4.9	Weitere Bestandteile von Beschreibungen . . . . .	110
4.5	Ableitung von Texten aus der DRS . . . . .	111
4.5.1	Grundsätzliche Überlegungen . . . . .	111
4.5.2	Text- und DRS-Aufbau . . . . .	113
4.5.3	Einfache Hauptsätze . . . . .	114
4.5.4	Relativsätze . . . . .	115

4.5.5	Prädikation . . . . .	116
4.5.6	Duplexbedingungen, 1. Teil . . . . .	116
4.5.7	Räumliche Beschreibungsergänzungen . . . . .	117
4.5.8	Duplexbedingungen, 2. Teil . . . . .	117
4.5.9	Satzfolgen basierend auf Duplexbedingungen . . . . .	118
4.5.10	Pronominalisierung . . . . .	119
4.5.11	Weitere Satzbautechniken . . . . .	123
4.6	DRS-Manipulation . . . . .	126
4.7	Lexika . . . . .	128
<b>5</b>	<b>Ergebnisse, Zusammenfassung und Ausblick</b>	<b>131</b>
5.1	Testszenarien . . . . .	131
5.2	Auswertungsergebnisse . . . . .	133
5.3	Zusammenfassung . . . . .	136
5.4	Bewertung . . . . .	137
5.5	Ausblick . . . . .	138
5.6	Andere Diskursbereiche . . . . .	143
<b>A</b>	<b>Auswertungsergebnisse</b>	<b>146</b>
A.1	Tankstellenszene . . . . .	146
A.1.1	Objekt 2 . . . . .	146
A.1.2	Objekt 1 . . . . .	149
A.1.3	Objekt 4 . . . . .	151
A.1.4	Objekt 10 . . . . .	153
A.1.5	Objekt 5 . . . . .	155
A.1.6	Beschreibung mit mehreren Hauptakteuren . . . . .	157
A.2	Alte Durlacher-Tor-Szene . . . . .	158
A.3	Ettlinger-Tor-Szene . . . . .	159
A.4	Nibelungenplatzszene . . . . .	161
A.5	Rheinhafenszene . . . . .	162
<b>B</b>	<b>Situationsgraphenbaum</b>	<b>165</b>
<b>C</b>	<b>MALCOLM</b>	<b>172</b>



C.1	Formatvorgaben . . . . .	172
C.2	Datenstrukturen . . . . .	174
C.3	Konstruktion des Perspektivierungsgraphen . . . . .	175
C.3.1	Initialisierung des Perspektivierungsgraphen . . . . .	175
C.3.2	Rekonstruktion der Nachfolgebeziehungen . . . . .	176
C.3.3	Rekonstruktion der Hierarchiebeziehungen . . . . .	177
C.4	Ausgabe des Perspektivierungsgraphen . . . . .	177
<b>D</b>	<b>ANGUS</b>	<b>178</b>
D.1	Übersicht . . . . .	178
D.2	Dateienorganisation . . . . .	179
D.2.1	/home/gerber/src/final . . . . .	179
D.2.2	/home/gerber/src/final/ <i>bildfolge</i> . . . . .	180
D.2.3	/home/gerber/src/limette . . . . .	180
D.2.4	/home/gerber/src/limette/sits . . . . .	181
D.3	Quelltext . . . . .	181
D.4	Lexika . . . . .	181
D.4.1	Verben und Hilfsverben . . . . .	181
D.4.2	Nomen . . . . .	182
D.4.3	Kardinalzahlen und Quantoren . . . . .	182
D.4.4	Personalpronomen . . . . .	183
D.4.5	Sonstige Wortarten . . . . .	183
D.5	Konstruktionsregeln und Grammatiken . . . . .	183
D.5.1	Grammatik . . . . .	184
D.5.2	Konstruktionsregeln . . . . .	186
D.6	Anbindung von LIMETTE . . . . .	187
D.6.1	limette_aufruf . . . . .	187
D.6.2	limette_perspektive . . . . .	188
D.6.3	limette_grupp . . . . .	189
	<b>Literaturverzeichnis</b>	<b>190</b>

# Abbildungsverzeichnis

1.1	Vier Einzelbilder aus der Tankstellenszene . . . . .	13
2.1	Modell für ‘Perspektive’ . . . . .	33
2.2	Systementwurf . . . . .	40
3.1	Allgemeiner Aufbau eines Situationsgraphenbaums . . . . .	48
3.2	Fahrbahnmodell der Tankstelle . . . . .	52
3.3	Perspektivische Beschreibungsvarianten . . . . .	56
3.4	Wirkungsszenario von MALCOLM . . . . .	58
3.5	Ableitung des Perspektivierungsbaumes . . . . .	59
3.6	Beispiel: Errechnete perspektivische Beschreibung . . . . .	61
3.7	Quantoren . . . . .	64
4.1	Syntaxanalyse von Fakten . . . . .	87
4.2	Strukturbaum zum Faktum <b>enter(object_4, kriegsstraße, 1)</b> . . . . .	91
4.3	Konstruktionsregel KR.EN (vorl.) . . . . .	92
4.4	DRS zum Faktum <b>enter(object_4, kriegsstraße, 1)</b> . . . . .	92
4.5	DRS zum Faktum <b>approach(object_4, ettlinger_straße, 2)</b> . . . . .	93
4.6	Konstruktionsregel KR.EN (endg.) . . . . .	94
4.7	Modifizierte DRS . . . . .	94
4.8	DRS für definite Beschreibungen . . . . .	95
4.9	Konstruktionsregel KR.DB . . . . .	96
4.10	Konstruktionsregel KR.PP . . . . .	97
4.11	DRS mit Prädikationen . . . . .	98
4.12	Konstruktionsregel KR.PRAED . . . . .	99
4.13	DRS für Quantoren und Kardinalzahlen . . . . .	99
4.14	Konstruktionsregel KR.Ballung . . . . .	100

4.15	DRS für Quantoren und Kardinalzahlen 2	102
4.16	DRS für Quantoren und Kardinalzahlen 3	103
4.17	Konstruktionsregeln für Quantoren	104
4.18	Angleichung von inneren DRSen	105
4.19	DRS für Kardinalzahlen 4	105
4.20	DRS für indefinite Beschreibungen	106
4.21	Konstruktionsregeln für indefinite Beschreibungen	107
4.22	Beispiel-DRS inkl. temporaler Aspekte (vorl.)	108
4.23	Zeitmodell nach [Allen 83]	109
4.24	Beispiel-DRS inkl. temporaler Aspekte (endg.)	110
4.25	Einfache Hauptsatzbildung	114
4.26	Relativsatzkonstruktion	115
4.27	Prädikation	116
4.28	Satzbauregel für Duplexbedingungen	117
4.29	Räumliche Beschreibungsergänzungen	118
4.30	Duplexbedingungen, 2. Teil	119
4.31	Duplexbedingungen und Satzfolgen	120
4.32	Beispiel-DRS	121
5.1	Straßenverkehrsszene am Ettlinger Tor.	132
5.2	Systementwurf	139
A.1	Objekt_2 der Tankstellenszene.	147
A.2	Fahrmanöver von Objekt_2	147
A.3	Fahrmanöver von Objekt_1	149
A.4	Objekt_1 der Tankstellenszene.	150
A.5	Fahrmanöver von Objekt_4	151
A.6	Trajektorie von Objekt_4 der Tankstellenszene (in der Momentaufnahme hinter dem Werbemast befindlich).	152
A.7	Objekt_10 der Tankstellenszene.	154
A.8	Fahrmanöver von Objekt_10	154
A.9	Objekt_5 der Tankstellenszene.	156
A.10	Fahrmanöver von Objekt_5	157
A.11	Durlacher-Tor-Szene	158

A.12 Ettliger–Tor–Szene . . . . .	160
A.13 Nibelungenplatzszene . . . . .	162
A.14 Rheinhafenszene . . . . .	163
C.1 Datenstrukturen in MALCOLM . . . . .	175

# Tabellenverzeichnis

2.1	Beispiel einer begrifflichen Beschreibung . . . . .	24
2.2	Beschreibungskonstellationen . . . . .	37
2.3	Anfragen für Beschreibungsvarianten . . . . .	38
3.1	Semantik von Logikoperatoren . . . . .	45
3.2	Ausdrücke in F-Limette . . . . .	46
4.1	Erläuterung zu Abbildung 4.1 . . . . .	88

# Kapitel 1

## Einleitung

### 1.1 Aufgabenstellung

Die vorliegende Arbeit beschäftigt sich mit der Generierung natürlichsprachlicher Beschreibungen zeitveränderlicher Szenen, die mit Hilfe einer stationären Videokamera aufgezeichnet wurden. Zu Beginn stellte dieses Problem reine Grundlagenforschung dar. Das bestehende, am Institut für Algorithmen und Kognitive Systeme (IAKS) entwickelte Bildfolgenauswertungssystem hatte einen hohen Reifegrad erreicht, so daß es große Mengen zuverlässiger Auswertungsdaten lieferte. Diese reichhaltige Fülle an Auswertungsdaten (in Form von Zahlentabellen oder Listen primitiv-begrifflicher Beschreibungen) machte deren Inspektion für den Benutzer zur zeitraubenden Arbeit. Somit wurde die Fragestellung interessant, wie dieses Datenmaterial in eine komprimierte Darstellung überführt werden konnte, welche sowohl eine effektive als auch effiziente Inspektion ermöglichte.

Die natürliche Sprache ist das Kommunikationsmittel, das der Mensch mit am häufigsten und intuitivsten benutzt. Deshalb erscheint das Bemühen, Ergebnisse eines Bildfolgenauswertungsprozesses als natürlichsprachlichen Text darzustellen, berechtigt zu sein. Zudem ist die natürliche Sprache ein sehr mächtiges Kommunikationsmittel. Sie erlaubt, Abläufe sehr detailliert in 'epischer Breite' zu beschreiben, oder aber auch sehr komprimiert in wenigen Sätzen. Verschiedene Beschreibungen derselben Szene können jedoch nicht nur in ihrem Abstraktionsgrad variieren. Statt dessen sind auch Unterschiede im Zweck der Beschreibung, ihrer Schwerpunktsetzung und in der Bewertung der beschriebenen Abläufe vorstellbar. Die natürlichsprachliche Beschreibung kann inkrementell oder a-posteriori unter Hinzuziehung erst nachträglich erworbenen, zur Zeit des eigentlichen Tathergangs noch nicht vorhandenen Wissens erstellt werden. Verbunden mit einem Antwortgenerierungssystem können gezielte Anfragen beantwortet werden, wobei nur die für die Antwort relevante Information versprochen wird. All diese Faktoren, die eine natürlichsprachliche Szenenbeschreibung beeinflussen, bilden zusammen die *Perspektive*, mit der die Szene betrachtet wird. Natürlichsprachli-

#340



#635



#1180

#2131

Abbildung 1.1: Vier Einzelbilder aus der Tankstellenszene inklusive Überlagerungen der automatisch berechneten Trajektorien und der angepaßten Fahrzeugmodelle für zwei erkannte Fahrzeuge der Szene (*obj\_4* und *obj\_9*). **#340**: *Obj\_9* fährt zum zweiten Tankplatz der oberen Tankspur. **#635**: *Obj\_4* überholt *obj\_9*. **#1180**: *Obj\_4* steuert den letzten Tankplatz an. **#2131**: *Obj\_9* überholt *obj\_4*, um selbst zur Tankstellenausfahrt zu gelangen.

che Beschreibungen sind also grundsätzlich perspektivisch. Durch das Perspektivische einer Beschreibung kann der Leser auf den individuellen Standpunkt des Erzählers rückschließen. Umgekehrt bestimmt der Standpunkt des Erzählers die Perspektive der Beschreibung.

Die anfänglichen Bemühungen, Bildauswertungsergebnisse natürlichsprachlich zu formulieren, trugen sehr bald erste Früchte. Es wurden Werkzeuge geschaffen, mit deren Hilfe formal und unter Minimierung von Heuristiken Daten abstrahiert werden konnten. Diese höherwertigen begrifflichen Beschreibungen wurden in einfache Texte

transformiert. Anhand dieser Texte konnte sehr schnell erkannt werden, ob bestimmte Auswertungsergebnisse plausibel waren oder nicht. Im letztgenannten Fall wurde Ursachenforschung betrieben, wodurch je nach Fehlerquelle die signalnahe Bildauswertung oder die begriffliche und natürlichsprachliche Ebene verbessert wurden. Im Laufe der Zeit haben sich die Anforderungen an die natürlichsprachliche Auswertung erhöht. Die signalnahe Bildfolgenauswertung ist hinreichend robust, um auch unter schwierigen Auswertungsbedingungen zuverlässige Auswertungsdaten zu liefern. Dies bildet die Grundlage, um vor allem die natürlichsprachliche Auswertung zu verfeinern und auszuweiten. Die Aufgabe der Textgenerierungskomponente soll sich fortan nicht mehr darauf beschränken, *irgendeine* natürlichsprachliche Beschreibung extrahieren zu können. Statt dessen muß das System in die Lage versetzt werden, abhängig von Benutzer- oder Spezifikationsvorgaben individuelle Erzählerstandpunkte zu definieren, bezüglich derer perspektivische Beschreibungen extrahiert werden können. In der vorliegenden Arbeit wird eine Methode entwickelt, die prinzipiell keine erlaubte Perspektive, mit der eine natürlichsprachliche Beschreibung erstellt werden kann, beschränkt. Um eine solche Methode realisieren zu können, empfiehlt es sich, natürlichsprachliche Beschreibungen, die von Menschen zu bestimmten Abläufen generiert wurden, auf Unterschiede und Gemeinsamkeiten zu untersuchen. Kann man die Einflußfaktoren ermitteln, die bei der Erstellung natürlichsprachlicher Beschreibungen eine Rolle spielen, so erscheint es prinzipiell möglich, Perspektivität generisch zu modellieren. Bei der Schaffung einer generischen Beschreibung von Perspektivität in Sprache und Text werden vorwiegend linguistische, literaturwissenschaftliche und philosophische Fragestellungen berührt. Die Realisierung perspektivischer natürlichsprachlicher Beschreibungen aus einer geometrischen oder primitiv-begrifflichen erfolgt in mehreren Stufen begrifflicher Abstraktion. Hierzu wird auf ein bestehendes unscharf metrisch-temporales Logikkalkül zurückgegriffen, wodurch in der vorliegenden Arbeit auch formal-logische Betrachtungsweisen relevant sind. Da die Arbeit zudem auf ein bestehendes Bildfolgenauswertungssystem aufbaut, mit diesem System interagiert und ausschließlich Realweltszenen behandelt, werden auch ingenieurwissenschaftliche Gesichtspunkte tangiert.

Als Diskursbereich wird in der vorliegenden und den zugehörigen Arbeiten am Institut der Bereich *Straßenverkehr* gewählt, da dieser Bereich einerseits ausreichende Komplexität besitzt, um verallgemeinerbare Einsichten zu ermöglichen, und andererseits ausreichende Einfachheit, um eine überschaubare Auswertung mit den zur Verfügung stehenden Mitteln zu ermöglichen ([Nagel 88]). Als Einsatzgebiet für ein System zur automatischen Generierung von natürlichsprachlichen Szenenbeschreibungen durch Bildfolgenauswertung kommen in diesem Diskursbereich Systeme zur statistischen Auswertung von Verkehrsflüssen an Straßen und Kreuzungen in Betracht. Auch an die Kommunikation des Menschen mit einem Fahrerassistenzsystem kann gedacht werden. Grundsätzlich besitzt der in der vorliegenden Arbeit entwickelte Ansatz für die Kommunikation aller Arten sichtsystembasierter Maschinen mit dem Menschen Relevanz. Darüberhinaus kann der Ansatz prinzipiell auf weitere Problembereiche angepaßt werden, so z.B. zur automatischen Archivierung von Videobändern.



## 1.2 Sprachverarbeitung in der Informatik

Die vorliegende Arbeit ist hauptsächlich dem Bereich der Bildfolgenauswertung zuzuordnen. In späteren Abschnitten wird auf das am Institut für Algorithmen und Kognitive Systeme entwickelte Bildfolgenauswertungssystem eingegangen. Allerdings schlägt die vorliegende Arbeit eine Brücke von der Bildfolgenauswertung zur maschinellen Sprachverarbeitung. Aus diesem Grund folgt eine kurze Einführung zu der Bedeutung natürlicher Sprache in der Informatik.

Bereits in den 60er Jahren gab es erste Anläufe, Rechner mit sprachlicher Kompetenz zu versehen. Diese anfänglichen Bemühungen haben sich bis heute in unterschiedliche Funktionalitäten entwickelt, die sich wie folgt darstellen lassen:

- **Textausgabe:** Der Rechner liest einen bereits in seinem Speicher vorhandenen Text vor.
- **Texteingabe:** Der Rechner überträgt einen natürlichsprachlichen Diskurs, der ihm über ein Eingabemedium (z.B. ein Mikrofon) zugeführt wird, in eine rechnerinterne Struktur.
- **Textmanipulation:** Der Rechner transformiert eine bereits vorhandene rechnerinterne Struktur, die einen Diskurs repräsentiert, in eine veränderte Struktur, beispielsweise zur Übersetzung von Texten von einer Sprache in eine andere, zur rechnergesteuerten Erstellung von Kurzfassungen oder zur Beantwortung von Anfragen.
- **Texterzeugung:** Der Rechner erzeugt selbst einen Text, d.h. baut eine entsprechende rechnerinterne Struktur auf. In diesem Text beschreibt der Rechner beispielsweise Ergebnisse sensorischer Aufzeichnungen, d.h. er stellt seine Interpretation einer von ihm beobachteten Szene in Form eines natürlichsprachlichen Textes dar.

Reine Texteingabesysteme haben bereits den Einzug in Privathaushalte gefunden. Man denke an Erweiterungen gängiger Textverarbeitungsprogramme, die es dem Benutzer erlauben, zu bearbeitende Texte in mündlicher Form zu diktieren, statt sie über eine Tastatur einzugeben. Solche reinen Texteingabesysteme erstellen jedoch in der Regel keine rechnerinterne Repräsentation des eingegebenen Textes, d.h. die Textsemantik bleibt ihnen verborgen. Üblicherweise sind Texteingabesysteme nicht ad hoc einsetzbar. Sie erfordern vorab genaue Abstimmung an den Sprecher und die Aufnahmeumgebung, oder beschränken den Wortschatz des Systems signifikant. Auch reine Textausgabesysteme sind bereits für den Privatkunden erhältlich. Prototypen von Übersetzungsprogrammen verlassen derzeit die Laborumgebung.

Letztgenannte Art der Programme berücksichtigen bereits die Semantik eines Textes, besitzen also eine höhere sprachliche Kompetenz. [Harris 85] bezeichnet *sprachverarbeitende Systeme* als Systeme, die zur Sprachanalyse, Sprachsynthese, Transformation

(einer sprachlichen Einheit in eine andere, äquivalente Einheit) und Schlußfolgerung fähig sind. Sprachverarbeitende Systeme gibt es in unterschiedlichen Ausprägungen. *Dialogsysteme* stellen kooperativ oder unkooperativ ([Jameson et al. 94]) in einer nachempfundenen Gesprächssituation Informationen bereit oder fordern diese an. *Interaktionssysteme* bewerkstelligen die Kommunikation zwischen Mensch und Maschine mittels natürlicher Sprache (vgl. [Lüth et al. 94]). *Auskunftssysteme* stellen dem Benutzer interaktiv Informationen zur Verfügung (z.B. Wegplanungssysteme, [Maaß 94]). Sprachkompetente *Auswertungssysteme* versprachlichen die sensorbasierte Interpretation einer Szene. Diese Aufstellung unterschiedliche Systemarten, die sprachliche Kompetenz nutzen, ist nicht vollständig. Es läßt sich auch nicht scharf zwischen den einzelnen hier aufgeführten Klassen trennen. Ein Bildfolgenauswertungssystem mit sprachlicher Kompetenz kann beispielsweise als Auswertungssystem aufgefaßt werden. Eingebunden in ein Fahrerassistenzsystem wird es jedoch auch zum Auskunftssystem. Da im Grunde jedes Programm Informationen (Daten) erzeugt, die für bestimmte Menschen von Interesse sein können, ziehen sich die Bemühungen, diese Informationen automatisch zu versprachlichen, durch unzählige Arten von Anwendungsgebieten, Spezifikationen und Diskursbereiche. Eine vergleichende Darstellung hierzu kann [Jones&Galliers 96] entnommen werden.

### 1.3 Literaturüberblick

Es folgen einige Ausführungen über auswärtige Arbeiten, die sich ebenfalls mit der Integration von Bildauswertung und natürlicher Sprache befassen, oder sonst in irgendeiner Weise mit der vorliegenden Arbeit vergleichbar sind.

Das System NAOS ([Neumann&Novak 86]) generiert natürlichsprachliche Beschreibungen von Straßenverkehrsszenen aus einer rein geometrischen Szenenbeschreibung (GSB). Die GSB enthält dabei idealerweise für jeden Zeitpunkt alle in der Szene sichtbaren Objekte, deren 3D-Positionen, Orientierungen, physikalische Oberflächeneigenschaften (z.B. Farbe) und deren Klassenzugehörigkeit (z.B. Automarke), sowie Angaben über den Betrachterstandpunkt und die Beleuchtungsdaten. Da den Autoren kein Bildauswertungssystem zur Erzeugung dieser angeforderten Daten zur Verfügung stand, wurde zu Experimentierzwecken ausschließlich mit synthetisch erzeugten GSBen gearbeitet. Um Aktivitäten in der Szene natürlichsprachlich beschreiben zu können, werden die in der GSB vorliegenden quantitativen Daten mit qualitativen Beschreibungsformen assoziiert. Dazu wird eine begriffliche Hierarchie von Ereignissen eingeführt, welche vorhandenes Wissen über relevante Bewegungen in der Szene durch generische Ereignismodelle repräsentiert.

[Fernyhough et al. 98] generieren Ereignismodelle anhand qualitativer Schlußfolgerungsmechanismen und einer statistischen Auswertung von Videodaten. Ihre Überlegungen wenden sie auf den Diskursbereich ‘Straßenverkehr’ an, beschränken sich jedoch auf die Auswertung von Szenen auf übersichtlichen Fernstraßen, welche weitaus geringere

Komplexität in den zu beobachtenden Abläufen aufweisen, als die im Rahmen dieser Arbeit untersuchten Kreuzungs- und Tankstellenszenen.

[Friedman 96] verwendet sogenannte ‘Lexikalische Begriffsstrukturen’ (LCSen), um geometrische Bildfolgenauswertungsergebnisse in eine natürlichsprachliche Beschreibung zu transformieren. Seine Überlegungen wurden jedoch nicht auf reale Bilddaten angewendet.

[Buxton&Gong 95], [Howarth&Buxton 98] stellen ein System zur sightssystembasierten Überwachung von Objektverhalten in Straßenverkehrsszenen vor. Die Autoren unterscheiden, analog zu unserem Ansatz, zwischen der signalnahen und der begrifflichen Auswertung. In erstgenannter werden geometrische Beschreibungen von Ergebnissen der Verfolgung sich bewegender Objekte zusammen mit dem angenommenen Objekttyp erzeugt. Zweitgenannte Auswertungsstufe konvertiert die geometrischen Beschreibungen in konsistente Verhaltensbeschreibungen, welche auf der Ereignis- und Verhaltensmodellierung sowie auf einem Kamera- und einem Lageplanmodell basieren. Als Repräsentationsstruktur, anhand derer die Verhaltensanalyse durchgeführt wird, verwenden die Autoren Bayes’sche Netzwerke zur Modellierung von Abhängigkeiten zwischen maßgeblichen Szenenparametern.

[Huang et al. 94] verwenden ebenfalls dynamische Wahrscheinlichkeitsnetzwerke, um Realwelt-Straßenverkehrsszenen zu interpretieren. Die Autoren beschränken sich auf die Auswertung von Autobahnsszenen, welche nicht annähernd die inhaltliche Komplexität der Szenen aufweisen, die von uns ausgewertet werden.

Wie die hier vorliegende Arbeit widmen sich auch [Dance et al. 96] der Auswertung von Realweltbildfolgen von Straßenverkehrsszenen und der Formulierung der Auswertungsergebnisse als natürlichsprachlichen Text. Auch in diesem Ansatz unterscheiden die Autoren zwischen signalnaher Bildauswertung und Abstraktionsschicht. Als Abstraktionsmechanismus wird ein logikbasiertes Verfahren (‘Network of Frames’) verwendet, welches die Behandlung von Unschärfe durch Verwendung der Dempster-Shafer-Theorie realisiert. Zu diesem Ansatz läßt sich kritisch anmerken, daß die Realisierung der signalnahen Bildauswertung im Vergleich zu unserem Ansatz eher unausgereift erscheint.

Die im Rahmen des Forschungsprojekts VITRA ([Herzog 92]) vorgenommenen Untersuchungen haben ihren Ursprung in dem Zusammenschluß eines bestehenden Systems zur Auswertung von Straßenverkehrsszenen mit einem unabhängig davon entstandenen Dialogsystem (vgl. [Nagel 88]). Daraufhin wurde besagtes Projekt auf unterschiedliche Ansätze zur Integration von natürlichsprachlichen Texten und Bildauswertung ausgeweitet. Dabei wird sowohl die Umwandlung von natürlichsprachlichen Aussagen in Bilder als auch der entgegengesetzte Ansatz betrachtet. Das Teilprojekt SOCCER ([André et al. 88]) generiert natürlichsprachliche Beschreibungen von Fußballspielszenen. Sein Kern besteht aus drei Funktionseinheiten. In der *Ereigniserkennungskomponente* werden zu der berechneten geometrischen Szenenbeschreibung Ereignisbeschreibungen assoziiert. Die *Selektionskomponente* wählt diejenigen Ereig-

nisse aus, welche verbalisiert werden sollen, und ordnet diese in eine Prioritätswarteschlange ein. Die *Sprachgenerierungskomponente* erstellt aus der in der Selektionskomponente zusammengestellten Datenmenge einen natürlichsprachlichen Text. Das System wird ergänzt durch die Subsysteme REPLAI ([Retz-Schmidt 92]) und ANTLIMA ([Schirra&Stopp 93], [Schirra 94]). Das Subsystem REPLAI berechnet aus den erkannten Ereignissen Intentionen und Pläne der jeweiligen Agenten, und führt diese ebenfalls der Selektionskomponente zu. Das Subsystem ANTLIMA überprüft die durch die Selektionskomponente ausgewählte Ereignismenge durch Konstruktion eines ‘mentalen Bildes’ auf ihre Aussagekraft und Korrektheit, und führt ihre Ergebnisse wiederum der Selektionskomponente zu. Zu den im Projekt VITRA verfolgten Ansätzen läßt sich kritisch anmerken, daß sowohl die Auswertung der Straßenverkehrsszenen als auch der Fußballspielszenen teils auf synthetischen Daten (insbesondere zur Erfassung des stationären Bildhintergrunds), teils auf realen Ergebnissen unserer Forschungsgruppe (von Anfang der 90er Jahre) beruhen, welche jedoch interaktiv zu einer umfassenderen geometrischen Szenenbeschreibung ergänzt wurden, insbesondere zur Vervollständigung lückenhafter Trajektorien.

Das System MOSES ermöglicht die Generierung inkrementeller Wegbeschreibungen in statischen 3D-Szenen als natürlichsprachlichen Text ([Maaß 94], [Gapp&Maaß 94]). Das System simuliert dabei einen Beifahrer, der dem Fahrer anhand einer Karte den Weg durch ein beiden unbekanntes Gebiet weist. Dabei produziert das System aus dem in der Karte gespeicherten Wissen (das Angaben über Abstände und Orientierungen miteinschließt) eine topographische Unterteilung des Wegs zum Ziel.

Das System KANTRA, ebenfalls ein VITRA-Projekt, stellt eine Schnittstelle für den bidirektionalen Mensch-Maschine-Dialog zur Robotersteuerung dar ([Lüth et al. 94]). Dieser Dialog soll in natürlicher Sprache unter Verwendung räumlicher Relationen (vgl. [Stopp et al. 94], [Gapp 94b], [Adorni 94], [Clay&Wilhelms 96]) erfolgen. Vier Interaktionsarten sind vorgesehen: *Aufgabenspezifikation* für den Roboter auf elementarer und abstrakter Ebene, *Ablaufprotokollierung* seitens des Roboters; *Fehlerprotokollierung* und *Aktualisierung* der Umgebungsbeschreibung seitens des Menschen. Die weiteren Ausführungen der Autoren beschränken sich auf die erste Interaktionsart. Das Dialogsystem läßt sich in drei Hauptmodule unterteilen: Die *Analysekomponente* überträgt die natürlichsprachliche Eingabe in eine Wissensstruktur. Die *Evaluierungskomponente* setzt die Eingabe in Arbeitsschritte, Rückfragen und Statusmeldungen um. Die *Generierungskomponente* übersetzt die auszugebende Information in natürliche Sprache.

[Intille 98] beschäftigt sich mit der Repräsentation und der Erkennung von Abläufen, an der mehr als ein Agens gemeinschaftlich beteiligt sind. Als Diskursbereich werden Spiele des Amerikanischen Fußballs gewählt, wobei versucht wird, die dort zu beobachtenden Spielzüge automatisch zu erkennen. Die Erkennung wird dadurch erschwert, daß der Diskursbereich sowohl hochgradig strukturiert als auch hochgradig variabel ist. Alle Spieler verfolgen zu Beginn eines Spielzuges bestimmte, vorher feststehende Pläne. Je nach Spielsituation sind die Spieler jedoch gezwungen, ihre Pläne zu modifizieren oder ganz aufzugeben. Diese situationsabhängigen Verhaltensänderun-

gen erschweren die Erkennung des ursprünglich zugrundeliegenden Planes, welche Voraussetzung zur Erkennung des durchgeführten Spielzuges ist. Als Eingabe erhält das System Trajektorien über alle sich bewegenden Objekte in der Szene. Die Trajektorien werden interaktiv aus dem Bildmaterial erstellt. Das resultierende Verhalten aller Objekte wird als Bündel parallel verlaufender Ereignisströme (pro Agens genau ein Strom) betrachtet. Diese Ereignisströme können zueinander bestimmten zeitlichen und kausalen Zusammenhängen unterliegen. Das System basiert auf dem Bayes'schen Netzwerkformalismus.

[Abella&Kender 96] entwickeln ein System zur automatischen Bestimmung und Beschreibung der Lage und Art von Nierensteinen in Röntgenbildern. Die Auswertung erfolgt anhand eines Einzelbildes, auf dem Rückgrat und Becken zu erkennen sein müssen. Anhand dieser Merkmale erfolgt die Anpassung eines Körpermodells an das Grauwertbild. Dazu werden mit Hilfe eines Kantendetektionsverfahrens die Abbildungen innerer Organe erkannt. Im nächsten Schritt werden durch ein Kreiserkennungsverfahren Nierensteine erkannt und deren Lage bestimmt. Zur Lagebestimmung werden alle systemintern modellierten räumlichen Relationen auf den Stein als zu lokalisierendes Objekt und auf alle erkannten Organe als Referenzobjekte angewandt. Redundante räumliche Beschreibungen werden anhand von Zusagegraden und Prioritätsvorgaben eliminiert. Ausgewählte Kombinationen verbliebener räumlicher Beschreibungen werden dann in die medizinischen Bezeichnungen der Lageorte übersetzt. Das Auswertungsergebnis wird als natürlichsprachlicher Text ausgegeben. Das System beschreibt keinerlei zeitliche Veränderungen in einer Szene, da es lediglich auf die Auswertung von Einzelbildern ausgerichtet ist.

Im Gegensatz zu den zuletzt vorgestellten Ansätzen widmet sich das von [Srihari 95] vorgestellte System PICTURE nicht der Ableitung natürlichsprachlicher oder begrifflicher Beschreibungen aus Bildfolgen, sondern dem entgegengesetzten Problem der Interpretation von Einzelbildern unter Rückgriff auf die Auswertung begleitender natürlichsprachlicher Texte (z.B. Bildunterschriften). Hierzu erstellt das System eine interne semantische Repräsentation des Textes, um dann durch Interpretation von Texthinweisen Objekte im Bild zu identifizieren.

Verbindungen zwischen Bild- und Sprachverarbeitung finden sich auch in Forschungsansätzen, die sich mit der Verwaltung von digitalen Videodatenbanken beschäftigen. [Bolle et al. 98] geben einen Überblick über aktuelle Forschungsrichtungen in diesem Bereich. Einige Arbeiten werden im folgenden kurz eingeführt.

[Sato et al. 97] entwickeln das System 'Name-It', welches Gesichter und Namen in Nachrichtenvideos zueinander assoziiert. Dieses System kann entweder zu einem gegebenen Gesicht den zugehörigen Namen ableiten, oder aber zu einem gegebenen Namen das zugehörige Gesicht im Bild lokalisieren. Der Name kann dabei entweder aus einer Bildeinblendung oder aus dem Begleitton entstammen. Das System beinhaltet eine Komponente, welche für jedes gegebene Wort die Wahrscheinlichkeit berechnet, daß es zu einem Gesicht korrespondiert. Diese 'natürlichsprachliche Schicht' umfaßt ein Wörterbuch, lexikalische und grammatikalische Analyse sowie Wissen um den gewöhn-

lichen Aufbau von typischen Nachrichtenvideos. Ähnlich zu [Srihari 95] widmen sich auch diese Autoren folglich dem - zu unserem Ansatz entgegengesetzten - Problem der Interpretation einer Bildfolge anhand begleitender linguistischer Information.

[Bloomberg&Chen 96] beschreiben ein System, welches Kurzfassungen von als digitalisierte Bilder vorliegenden Textdokumenten erstellt. Es werden, basierend auf einer statistischen Auswertung des Auftretens charakteristischer Worte, zusammenfassende Sätze extrahiert. Dazu wird das Bild in Textregionen, Zeilen und Worte segmentiert, sowie Sätze und Abschnittsgrenzen erkannt. Darauf aufbauend wird eine Menge von Wortäquivalenzklassen berechnet. Die Auftrittshäufigkeit eines Worts, dessen Stellung im Satz und dessen Wortlänge werden herangezogen, um die Äquivalenzklassen hinsichtlich ihrer Bedeutung zu bewerten. Darauf aufbauend wird jeder Satz hinsichtlich seiner Wichtigkeit beurteilt. Die  $n$  am besten bewerteten Sätze bilden die Zusammenfassung des Textdokuments. Auch wenn [Bloomberg&Chen 96] natürlichsprachlichen Text durch Bildauswertung erzeugen, so bauen sie doch keine interne Repräsentation der Semantik des Textes auf. Daraus folgt, daß das System nicht in der Lage ist, eigenständige Sätze zu formulieren.

Einen im Grundsatz ähnlichen Ansatz verfolgen [Smith&Kanade 97], welche Zusammenfassungen von Videos erstellen, indem sie automatisch Bereiche erkennen, welche entweder wichtige Ton- oder Bildinformation beinhalten. Die relative Wichtigkeit jeder Szene wird anhand der im Bild auftretenden Objekte, der zugehörigen Worte und der Struktur der Szene bestimmt. Die Sprachanalyse identifiziert wichtige Tonregionen durch Berechnung der relativen Häufigkeit des Auftretens einzelner Worte. Worte, die statistisch gesehen im Vergleich zu Standardtexten relativ häufig auftreten, werden zu Schlüsselwörtern. Ein erster Auswertungsschritt liefert als Ergebnis eine reduzierte Tonspur, welche im wesentlichen nur aus diesen Schlüsselwörtern besteht. Um die Verständlichkeit des Textes zu erhöhen, werden die Schlüsselwörter zu 'Schlüsselphrasen' erweitert, welche auch die 'Umgebung' der Schlüsselwörter enthalten. Auch dieser Ansatz ermöglicht keine interne Repräsentation der Semantik des Textes, so daß analog zu [Bloomberg&Chen 96] kein eigenständiger, grammatikalisch korrekter Text aufgebaut werden kann.

[Courtney 97] erkennt Bewegungen aus Videobildfolgen und repräsentiert diese mit gerichteten Graphen, die das jeweilige Objekt und die jeweilige Bewegung beschreiben. Die Aufspürung bestimmter Abläufe in der Videobildfolge erfolgt dann nicht mehr durch Inspektion der Rohdaten, sondern durch Inspektion des Bewegungsgraphen.

[Socher et al. 98] verbinden Bildauswertung und Sprache, indem sie in einer 'Bauklötzchenwelt' bestimmte Objekte aufspüren, die vom Benutzer natürlichsprachlich beschrieben werden (z.B. 'finde blauen Würfel' etc.). Das System umfaßt die Analyse der Anfrage, die Extraktion der geforderten Objektmerkmale und die Berechnung räumlicher Relationen aus der mit einer stationären Kamera aufgezeichneten Bildfolge. Das System basiert auf Bayes'schen Netzwerken und betrachtet nicht die Beschreibung oder das Aufspüren von Bewegungen in einer Szene. Einen vergleichbaren Problembereich bearbeiten [Chella et al. 97], allerdings unter der Verwendung terminologischer Systeme.

me.

Das Projekt PRACMA ([Jameson et al. 94]) beschäftigt sich mit der Entwicklung eines Dialogsystems. Das System wird auf dem Diskursbereich ‘Gebrauchtwagenkauf’ aufgesetzt, wobei das System einerseits den gewinnorientierten Verkäufer, andererseits den mißtrauischen Kunden simulieren kann. Es strebt also die algorithmische Behandlung *unkooperativer* Dialoge an, bei denen die Gesprächspartner unterschiedliche Ziele verfolgen. Das Projekt PRACMA beschäftigt sich nicht mit der Verarbeitung visueller Daten, jedoch in hohem Maße mit der rechnerinternen Analyse und Synthese natürlicher Texten. Dabei werden bevorzugt die folgenden Fragestellungen untersucht: Wie kann ein Informationssuchender Informationen aus einem unkooperativen Informationshalter herauslocken und interpretieren, und wie kann er mit den in den Informationen enthaltenen Unsicherheiten umgehen? Wie kann ein Informationsanbieter die Schlußfolgerungen und Überlegungen eines größtenteils unbekanntes Informationssuchenden rekonstruieren, vorhersehen und manipulieren?

Auf die einzelnen Arbeiten wird im Verlauf der vorliegenden Arbeit immer dann, wenn es angebracht erscheint, ausführlicher eingegangen.

# Kapitel 2

## Natürlichsprachliche Beschreibungen

### 2.1 Inhalt und Stil

Natürliche Sprache als Kommunikationsform im Dialog von Mensch und Maschine kann nur dann ihren Zweck erfüllen, wenn zum einen die maschinell versprachlichte inhaltliche Information ausreichend und zuverlässig ist. Zum anderen muß eine maschinell erzeugte natürlichsprachliche Szenenbeschreibung in einer Form präsentiert werden, die vom Menschen als adäquat angesehen wird. Mit anderen Worten: Die Güte eines (insbesondere maschinell erstellten) Textes hängt nicht allein von seinem Inhalt, sondern ebenso von seiner Form ab. Die inhaltliche Information muß nicht nur zutreffend sein, sie muß gleichzeitig auch gut strukturiert, flüssig, prägnant und damit möglichst frei von Mehrdeutigkeiten und Redundanz präsentiert werden. Bereits hier erweisen sich die Begriffe *Inhalt* und *Stil* sowie die Trennung dieser beiden Begriffe als essentiell bei der maschinellen Generierung (nach den genannten Bewertungskriterien) angemessener natürlichsprachlicher Szenenbeschreibungen. Grundsätzlich wird dabei der Inhalt zuerst betrachtet. Erst wenn feststeht, was ausgesagt werden soll, wird im nächsten Schritt evaluiert, wie diese Aussage gestaltet wird. Das Bildfolgenauswertungssystem wird folglich um zwei Systemschichten ergänzt. Die *begriffliche Schicht* extrahiert die inhaltliche Information und leitet diese an die *natürlichsprachliche Schicht* weiter. Diese Schicht bestimmt dann den stilistischen Aufbau der entstehenden Beschreibung. In späteren Betrachtungen wird sich jedoch zeigen, daß nicht nur der Stil vom Inhalt abhängig ist. Vielmehr wird sich erweisen, daß stilistische Beschränkungen der natürlichsprachlichen Schicht Einfluß auf die Ausgestaltung der begrifflichen Schicht haben werden und Wechselwirkungen zwischen beiden Systemebenen notwendig sind. Es wurde bereits festgestellt, daß die Güte einer maschinell generierten natürlichsprachlichen Beschreibung davon abhängt, inwieweit den Anforderungen des menschlichen Empfängers der Beschreibung Rechnung getragen werden konnte. Es empfiehlt sich



daher, bei der Konstruktion eines Systems zur maschinellen Generierung natürlicher Beschreibungen den Sprachgebrauch des Menschen als Vorbild zu nehmen. Die folgenden Überlegungen sollen einige der dabei relevanten Fragestellungen illustrieren. Detailliertere Betrachtungen, wie besagte Fragestellungen gelöst werden, erfolgen dabei später. Es wird eine natürlichsprachliche Beschreibung einer Szene angegeben, die sich an einer Tankstelle abgespielt hat. Diese Beschreibung wurde von einem Menschen erstellt ([Kamp&Schiehlen 99], [Gerber 99]). Sie ist Bestandteil eines kritischen Vergleichs zwischen systemintern erstellten natürlichsprachlichen Beschreibungen und natürlichsprachlichen Beschreibungen, wie sie von Menschen erzeugt werden. Das Beispiel soll die Zweckmäßigkeit der Trennung von Inhalt und Stil noch einmal bekräftigen. Es bezieht sich auf die Tankstellenszene aus Abbildung 1.1 und beschreibt das Verhalten des Fahrzeugs ‘Objekt\_4’. Die einzelnen Sätze sind zur späteren Referenzierung mit römischen Zahlen gekennzeichnet.

*(I) Object\_4 entered the filling station. (II) On both filling lanes the second filling place was occupied. (III) Object\_4 chose the upper lane and approached the first filling place. (IV) Then it changed to the passing lane, overtook object\_9 and went to the pump in front of that vehicle. (V) When it got to this filling place, it did not stop, but continued to a position in front of it and then backed into it. (VI) There it came to a stop and took petrol.*

Die Beschreibung ist in englischer Sprache formuliert, da im Rahmen dieser Arbeit nur englischsprachige Texte erzeugt werden. Dies liegt zum einen daran, daß der zur rechnerinternen semantischen Repräsentation verwendete Formalismus ([Kamp&Reyle 93]) auf die Behandlung englischsprachiger Texte zugeschnitten wurde. Zum anderen sind englischsprachige Texte von ihrer Satzstruktur im allgemeinen ‘einfacher’ aufgebaut als beispielsweise deutschsprachige Texte. Dadurch wird zum Zwecke der vorläufigen Problemvereinfachung die Ableitung von Texten aus einer rechnerinternen Repräsentation erleichtert. Prinzipiell erscheinen die im Rahmen dieser Arbeit entwickelten Methoden hinreichend, um die Modalitäten und Besonderheiten anderer Sprachen ebenso repräsentieren zu können. Demzufolge erscheinen mittelfristig auch Ausgaben in anderen Sprachen möglich.

Zur Veranschaulichung der inhaltlichen und stilistischen Bestandteile in der oben angegebenen Beschreibung wird deren inhaltliche Information in Tabelle 2.1 durch Logikprädikate illustriert ([Gerber 99]). Die erste Spalte gibt jeweils den Zusicherungsgrad  $\lambda \in [0, 0; 1, 0]$  jedes Faktums an. Da eine natürlichsprachliche Beschreibung ohne linguistische Hecken angegeben ist, sind alle Fakten mit höchstmöglichem Zusicherungsgrad versehen. Der Angabe des Zusicherungsgrads folgt, abgetrennt durch einen senkrechten Strich, die Angabe des Gültigkeitszeitraums der angegebenen Relationen. Er gibt den zeitlichen Verlauf des beschriebenen Ablaufs als linear geordnete Zeitpunktmenge wieder. Die angegebenen Quantitäten sind willkürlich gewählt, wesentlich sind die qualitativen Werte, also das Auftreten zeitlicher Überlappungen oder Überschneidungen. Abgetrennt durch ein Ausrufezeichen folgen der Relationsnamen und die Argumente der Relation. Jede Relation enthält als letztes Argument einen üblicherweise ganzzahligen

---

1	614 :	614 !	enter(object_4, the_filling_station, 1).
1	616 :	616 !	be_occupied_on(the_second_filling_place, the_lower_filling_lane, 2).
1	618 :	618 !	be_occupied_on(the_second_filling_place, the_upper_filling_lane, 3).
1	619 :	619 !	choose(object_4, the_upper_filling_lane, 4).
1	621 :	626 !	approach(object_4, the_first_filling_place, 5).
1	627 :	881 !	change_to(object_4, the_passing_lane, 6).
1	627 :	881 !	overtake(object_4, object_9, 7).
1	627 :	881 !	go_to(object_4, the_last_filling_place, 8).
1	883 :	888 !	reach(object_4, the_last_filling_place, 9).
1	889 :	983 !	leave(object_4, the_last_filling_place, 10).
1	1081 :	1165 !	back_into(object_4, the_last_filling_place, 11).
1	1227 :	4707 !	stop_on(object_4, the_last_filling_place, 12).
1	1227 :	4707 !	take(object_4, petrol, 13).

---

Tabelle 2.1: Beispiel einer begrifflichen Beschreibung, die im wesentlichen dieselbe inhaltliche Information enthält wie die natürlichsprachliche Beschreibung I - VI.

gen Markierer **MK**, mit dem das entsprechende Faktum referenziert wird. Diese Referenzierung von Fakten wird benötigt, um beispielsweise kausale Beziehungen zwischen Fakten beschreiben zu können. Ausdrücke in der eben beschriebenen Notation werden fortan stets als *unscharf metrisch temporale Logikfakten* aufgefaßt und bezeichnet.

Oben angegebene Logikfakten geben eine erste Vorstellung davon, welche Teile einer natürlichsprachlichen Beschreibung zum Inhalt dieser Beschreibung gerechnet werden. Dies ist insbesondere die Benennung der einzelnen Teilabläufe, hier durch Relationen mit einer Verbform als Relationsnamen. Desweiteren ist die Evaluierung der zeitlichen Bezüge der einzelnen Teilabläufe untereinander Aufgabenbestandteil der inhaltsermittelnden Schicht, hier geschehen durch die Angabe der Gültigkeitsintervalle. Desweiteren sei die Beschreibung von kausalen Beziehungen als Aufgabe dieser Schicht noch erwähnt.

Aus der angegebenen **begrifflichen** Beschreibung läßt sich unmittelbar eine natürlichsprachliche Beschreibung ableiten. Das erste Argument jeder Relation läßt sich als Subjekt eines Satzes auffassen, das Prädikat des jeweiligen Satzes wird aus der Verbalphrase des Relationsnamens gebildet. Das jeweilige zweite Argument jeder Relation kann, sofern es sich dabei nicht um den Faktenmarkierer handelt, als objektive Satzergänzung aufgefaßt werden. Anhand dieser Satzbauregel werden zur natürlichsprachlichen Präsentation der in der begrifflichen Beschreibung enthaltenen Information lediglich zwei stilistische Entscheidungen getroffen: die Wahl der Zeitform (hier: 'Past Tense') und der Reihenfolge der Faktenversprachlichung (hier: streng chronologisch).

In der Tat enthält die oben angegebene natürlichsprachliche Beschreibung I - VI stilisti-

sche Entscheidungen, die weit über dieses Maß hinaus gehen. Diese sollen im folgenden diskutiert werden. Satz I entspricht der Versprachlichung des Faktums 1 aus Tabelle 2.1. Er wurde nach der zuvor geschilderten Satzbauregel gebildet. Satz II faßt die Fakten 2 und 3 der Tabelle zusammen. Bereits die Zusammenfassung dieser beiden Fakten zu einem Satz ist eine stilistische Entscheidung, die dem Menschen zwar naheliegend erscheint, dem System aber erst einmal ermöglicht werden muß. Auch das Paraphrasieren der ‘oberen und unteren Tankspur’ zu dem Ausdruck ‘beide Tankspuren’ schafft keine neuen inhaltlichen Erkenntnisse, sondern ist eine stilistische Entscheidung. Sie wird erst dann möglich, wenn spätestens bei der Textplanung berücksichtigt wird, daß generell nur zwei Tankspuren in der Szene vorhanden sind, so daß durch die Verwendung des Quantors ‘*both*’ keine Mißinterpretationsmöglichkeiten entstehen. Dies stellt ein Basiswissen dar, das rechnerintern repräsentiert sein muß und nicht explizites Bestandteil der begrifflichen Beschreibung aus Tabelle 2.1 ist. Satz III versprachlicht die Fakten 4 und 5 der Tabelle. Hier erschien es sinnvoll, beide Fakten in einem Satz zusammen zu fassen. Auch dies ist eine rein stilistische Entscheidung. Satz IV versprachlicht die Fakten 6-8. Stilistische Entscheidungen stellen hier die Pronominalisierung des Agens der Beschreibung und die Assoziation des Temporaladverbs ‘*then*’ zur zeitlichen Verankerung des neuen Satzes an seinen Vorgänger dar. Die Paraphrasierung des ‘letzten Tankplatzes’ zur ‘Zapfsäule vor jenem Fahrzeug’ ist in erster Linie stilistisch motiviert. Sie enthält zwar auch inhaltliche Information, wie die Zuordnung von ‘Object\_9’ zur Kategorie ‘Fahrzeug’, die Assoziation des Begriffs ‘Tankplatz’ zum Begriff ‘Zapfsäule’ und die explizite Beschreibung der räumlichen Relation zwischen Patiens und letztem Tankplatz. All dies stellt jedoch keine neue Information dar. Statt dessen handelt es sich um ableitbares Wissen, das aus der begrifflichen Beschreibung durch Hinzuziehung von Kontextwissen extrahiert werden kann. Satz V stellt eine sehr starke stilistische Variation der in den Fakten 9-12 enthaltenen Information dar. Besonders die Nebensatz-/Hauptsatzkonstruktion und die explizite Erwähnung des Scheiterns einer Handlungserwartung ‘*did not stop*’ sind im Kern stilistische Entscheidungen. In Satz VI wird mit der adverbialen Bestimmung ‘*there*’ auf einen Ort verwiesen, der sich aus dem bisherigen Beschreibungskontext ergibt. Solche Formulierungsentscheidungen zu ermöglichen und zu verwenden stellt auch eine Aufgabe der Textstilistik dar.

Die vorangegangenen exemplarischen Überlegungen sollten ein Gefühl dafür entwickeln helfen, welche Aspekte natürlichsprachlicher Beschreibungen deren Inhalt bzw. deren Stil zugeordnet werden. Inhalt und Stil werden im folgenden, soweit möglich, getrennt voneinander betrachtet. Die vorliegende Arbeit behandelt zum einen die Extraktion begrifflicher Beschreibungen durch Bildauswertung (also die Inhaltsentwicklung) wie auch die rechnerinterne Repräsentation begrifflicher Beschreibung. Diese Repräsentation liefert die Grundlage zur Betrachtung beschreibungstilistischer Fragestellungen in einer eigenen Systemschicht.

## 2.2 Erzählstil

Es ist einleuchtend, daß ein bestimmter Ablauf in einer Szene auf mehr als eine Weise beschrieben werden kann. Es ist sogar zu erwarten, daß derselbe Ablauf von verschiedenen Erzählern unterschiedlich interpretiert und beschrieben wird. Sich widersprechende Zeugenaussagen zu einem Tathergang seien hierbei als Beispiel genannt. Diese Unterscheidungen ergeben sich aus den individuellen, nicht nur räumlich zu verstehenden Standpunkten der jeweiligen Erzähler. Auf den Einfluß und die Festlegung einzelner Erzählercharakteristiken wird an späterer Stelle eingegangen. In diesem Kapitel soll der Begriff 'Erzählstil' beleuchtet werden, der ebenfalls den Wortlaut einer Beschreibung beeinflusst. Wie der Begriff schon aussagt, ist die Festlegung des Erzählstils eine Aufgabe der stilbestimmenden Auswertungsschicht.

Im Bereich der literaturwissenschaftlichen Erzähltheorie werden hauptsächlich Erzähltechniken in fiktionalen Texten (Romane etc.) untersucht. Da die dabei entwickelten Erkenntnisse und Theorien auf die hier vorliegende Forschungsarbeit nicht unmittelbar übertragbar sind, wird hierauf nicht näher eingegangen. Darüberhinaus lassen sich nur vergleichsweise wenige Arbeiten über Erzähltechniken in der Alltagssprache finden.

[Quasthoff 87] analysiert Erzählungen aus dem kommunikativen Alltag. Die Autorin stellt fest, daß derselbe Ablauf von derselben Person in unterschiedlichen Gesprächssituationen ganz unterschiedlich erzählt werden kann. So hat beispielsweise der Bericht über den Hergang eines Unfalls bei der polizeilichen Vernehmung eine andere sprachliche Form als blumige Schilderungen des Hergangs am Krankenbett, auch wenn inhaltlich im wesentlichen dieselbe Information wiedergegeben wird. Die linguistische Narrativik unterscheidet narrative Erzählungsformen in mehrererlei Hinsicht (vgl. [Quasthoff 87]):

- Zum einen wird zwischen dem *sachlichen Bericht* und der *szenischen Erzählung* unterschieden. Dabei wurde letztgenannter Bereich bisher eindringlicher erforscht, wobei diese Forschungsausrichtung meist stark in soziolinguistische Bereiche einmündete (d.h. Untersuchungen zur Abhängigkeit szenischer Erzählmerkmale von z.B. Schicht und Geschlecht). Ein *Bericht* stellt die Rückschau auf etwas Vergangenes aus der Sicht der Gegenwart des Sprechvorgangs dar. Bei der *szenischen Erzählung* hingegen übernimmt der Erzähler die Perspektive eines aktiv oder passiv am zurückliegenden Ereignis Teilhabenden.
- Eine andere Unterscheidung ergibt sich aus der Person des Erzählenden. Erzählt werden kann zum einen a posteriori aus der Sicht eines *allwissenden Erzählers*. Dessen Schilderung kann Wissen enthalten, das erst zum Erzählzeitraum, nicht aber bereits zum Zeitraum des tatsächlichen Ablaufs, vorhanden ist. Erzählt werden kann zum anderen aus der Sicht eines der *Protagonisten* und drittens aus der Sicht eines *Beobachters*.
- Eine dritte hilfreiche Unterscheidungsheuristik ergibt sich für die Autorin aus

dem Konzept des *Planbruchs* in einer Erzählung. Dieses Konzept geht, vereinfacht ausgedrückt, davon aus, daß das Erzählte deshalb erzählt wird, weil sich etwas ‘Außergewöhnliches’ ereignet hat. Dieses außergewöhnliche Ereignis stellt einen Planbruch dar. Erzählungen lassen sich somit auch dadurch unterscheiden, ob sich dieser Planbruch im Handlungsplan eines *Agenten* (Agentenplanbruch ‘APB’) oder aber in den Erwartungen eines *Beobachters* (Beobachterplanbruch ‘BPB’) manifestiert.

Aus dem Vorhandensein dreier möglicher *Erzählertypen* (Protagonist *Prot*, allwissender Erzähler *Erz*, Beobachter *Beob*), zweier semantischer Typen (APB und BPB) und zweier Diskursmuster (szenische Erzählung *SE* und Bericht *Ber*) ergeben sich rein kombinatorisch zwölf Variationen narrativer Strukturen. Davon werden acht von der Autorin wegen logischer Unverträglichkeit ausgeschlossen. Es bleiben folgende vier mögliche Kombinationen übrig:

1. Prot/APB/SE,
2. Beob/BPB/SE,
3. Erz/APB/Ber,
4. Erz/BPB/Ber.

Empirische Untersuchungen zeigen, daß die dritte Möglichkeit Erz/APB/Ber erzähltechnisch zwar möglich ist, real jedoch nicht oder nur selten auftritt. Wird von einem Agentenplanbruch erzählt, so nimmt der Erzählende gewöhnlich nicht die Perspektive eines neutralen Erzählers ein, sondern erzählt aus der Perspektive des betroffenen Protagonisten, wobei zusätzlich nicht das Diskursmuster ‘Bericht’, sondern das Erzählmuster ‘Szenische Erzählung’ gewählt wird (Konstellation 1: Prot/APB/SE). Zusammenfassend lassen sich also folgende drei Aussagen treffen:

- Wird von einem Agentenplanbruch erzählt, so nimmt der Erzählende die Perspektive eines unmittelbar Beteiligten des Ablaufs ein. Auf diese Weise entsteht eine **szenische Erzählung aus Protagonistensicht**.
- Wird von einem Beobachterplanbruch erzählt, so gibt es zwei Möglichkeiten. Zum einen kann eine **szenische Erzählung aus Beobachtersicht** entstehen.
- Andererseits kann der Erzählende den Beobachterplanbruch im Nachhinein aus der Sicht eines allwissenden Erzählers schildern. Dadurch entsteht ein **Bericht**.

Bei der Übertragung dieser Erkenntnisse auf die hier vorliegende Forschungsarbeit ist mehrererlei zu beachten. Zum einen betreffen die Ausführungen aus [Quasthoff 87] gesprochene Erzählungen, während in der hier vorliegenden Arbeit geschriebene Erzählungen betrachtet werden. Die Unterschiede, die sich daraus ergeben, sind jedoch nicht einschneidend. So werden durch die Entscheidung, eine Erzählung nicht auszusprechen,

sondern aufzuschreiben, keine neuen Erzählstile geschaffen. Es ändert sich lediglich die sprachliche Qualität der Beschreibung. So wird bei der Niederschrift beispielsweise mehr auf korrekte Einhaltung grammatikalischer Regeln und die Erzeugung vollständiger Sätze geachtet als bei der mündlichen Erzählung. Dennoch wird im folgenden der Umstand, daß Texte in ‘Schriftsprache’ und nicht in gesprochener Sprache erzeugt werden, besonders betont, indem fortan statt des Begriffs ‘szenische Erzählung’ der Begriff ‘szenische Beschreibung’ verwendet wird.

Desweiteren sind bei der Übertragung obiger Erkenntnisse noch zwei Begriffsklärungen vorzunehmen. Sowohl in [Quasthoff 87] als auch in der hier vorliegenden Arbeit wurde der Begriff *Erzähler* als Bezeichnung für den Urheber einer natürlichsprachlichen Beschreibung verwendet. Bei der maschinellen Generierung natürlichsprachlicher Beschreibung ist dieser Urheber eine Maschine. Dies würde bedeuten, daß nur genau ein Erzähler existiert, der ein festgelegtes Profil besitzt. Damit könnte jeder Ablauf nur vom Standpunkt dieses Erzählers beschrieben werden. Statt dessen wird jedoch angestrebt, Abläufe von unterschiedlichen Standpunkten aus zu beschreiben. Dies bedeutet, daß rechnerintern ‘virtuelle Erzähler’ geschaffen werden, von denen jeder ein eigenes Profil besitzt. Diese virtuellen Erzähler können generisch beschrieben und durch Benutzervorgaben ausgeprägt werden. Auf diese Problematik wird später eingegangen.

[Quasthoff 87] verwendet den Begriff ‘Protagonistensicht’, wenn sie über Erzählungen referiert, die bestimmte Handlungen aus der Sicht des entsprechenden Handlungsträgers schildern. In der hier vorliegenden Arbeit werden Handlungsträger generell als ‘Agenten’ bezeichnet. Dennoch lassen sich die Begriffe ‘Protagonist’ und ‘Agens’ nicht gleichsetzen. So können in einer natürlichsprachlichen Beschreibung mehrere Agenten auftreten, aber stets nur ein Erzähler. Deshalb wird gegebenenfalls nur dasjenige Agens, das gleichzeitig auch Erzähler ist, als Protagonist bezeichnet.

Als Ergebnis der in diesem Abschnitt vorgenommenen Untersuchungen ist festzuhalten, daß drei mögliche Erzählstile existieren, in denen natürlichsprachliche Beschreibungen formuliert werden können. Zur Verdeutlichung der Unterschiede dieser drei Erzählstile *Bericht*, *szenische Beschreibung aus Protagonistensicht* und *szenische Beschreibung aus Beobachtersicht* werden im folgenden beispielhaft drei maschinell erstellte natürlichsprachliche Beschreibungen angegeben, wobei jede einen anderen der drei Erzählstile wiedergibt. Alle drei Beschreibungen beschreiben das Verhalten desselben Fahrzeugs, nämlich des Fahrzeugs ‘object\_4’ aus der Tankstellenszene (Abbildung 1.1). Die drei Beschreibungen lauten wie folgt:

1. **Bericht:** Object\_4 entered the filling station. On the lower filling lane the second filling place was occupied by object\_2. On the upper filling lane the second filling place was occupied by object\_9. Object\_4 chose the upper filling lane. It headed for the last filling place. Therefore, it had to overtake object\_9. Object\_4 changed to the passing lane. It passed object\_9. Then it changed back to the filling lane. Then it backed up to the last filling place. Then it took in petrol. Later, object\_4 left the filling station.

2. **Szenische Beschreibung aus Beobachtersicht:** Object\_4 entered the filling station. It drove on the upper filling lane. Then it changed to the passing lane. It overtook another car standing on the second filling place. Then it changed back to the filling lane. It stopped behind the last filling place. Then it backed up to that filling place. It stopped. Then it took in petrol. After that, it left the filling station.
3. **Szenische Beschreibung aus Protagonistensicht** I needed petrol. So I entered the filling station. I saw that on the lower filling lane the second filling place was occupied. On the upper filling lane, the second filling place was occupied, too. I chose the upper filling lane and headed for the last filling place. Therefore, I had to overtake the car standing on the second filling place. I changed to the passing lane. I passed the car. Then I changed back to the filling lane. I stopped behind the last filling place. So I had to back up to the last filling place. Finally, I stopped on it. Then I took in petrol. After that, I left the filling station.

Am augenfälligsten sind die Unterschiede der szenischen Beschreibung aus Protagonistensicht zu den anderen beiden Beschreibungen. Da der Protagonist Agent und Erzähler in einem ist, beschreibt er seine persönlichen Wahrnehmungen üblicherweise in der ersten Person. Er kennt naturgemäß nur seine Sicht einer bestimmten Situation. Daraus folgt, daß er sein Verhalten detailliert begründen kann und wird. Über die Hintergründe des Verhaltens anderer beteiligter Agenten kann er jedoch nur Vermutungen treffen. Der Protagonist wird geneigt sein, fremdes Verhalten zu beobachten, zu beschreiben, zu bewerten, sein eigenes Handeln danach auszurichten und dementsprechend zu begründen und zu beschreiben. Szenische Beschreibungen aus Protagonistensicht geben den Ablauf, wie bereits ausgeführt, streng chronologisch wieder, unabhängig davon, ob der Erzählzeitraum mit dem Ablaufzeitraum übereinstimmt oder nicht.

Letztgenanntes gilt auch für szenische Beschreibungen aus Beobachtersicht (zweites Beispiel). Beobachtet ein Beobachter mehr als einen Agenten, so ist er frei in der Bestimmung einer Untermenge dieser Agenten, die für seine Beschreibung relevant sind. Diese beschreibungsrelevanten Agenten werden im folgenden als *Hauptakteure* bezeichnet. Es sei an dieser Stelle darauf hingewiesen, daß es bei der szenischen Beschreibung aus *Protagonistensicht* nur einen Hauptakteur gibt, der zudem auch noch fix ist: der Protagonist selber. Das Verhalten anderer Agenten spielt in einer solchen Beschreibung zwar auch eine wesentliche Rolle, wird aber grundsätzlich nur relativ zum Verhalten des Hauptakteurs betrachtet.

Wird davon ausgegangen, daß es sich um keinen zufälligen Beobachter handelt, so wird die Wahl der Hauptakteure möglicherweise durch die Funktion des Beobachters bestimmt. Wird der (maschinelle) Beobachter beispielsweise abstrakt als Teil einer Verkehrszählungseinheit betrachtet, so sind all diejenigen Fahrzeuge, die auf einer bestimmten Straße in eine bestimmte Richtung fahren, Hauptakteure. Da der Beobachter sich außerhalb des betrachteten Ablaufs befindet, d.h. selbst nicht in den Ablauf eingreift, wird er geneigt sein, das Verhalten der beobachteten Agenten nach seinen

eigenen Maßstäben zu bewerten, zu beschreiben und Mutmaßungen über die weitere Entwicklung des Ablaufs anzustellen, sofern seine Funktion dies erlaubt.

Der Berichterstatter in einem Bericht befindet sich ebenfalls außerhalb des Ablaufs, muß diesen jedoch noch nicht einmal selbst beobachtet haben. Man denke beispielsweise an einen Polizeibericht über einen Unfall, der üblicherweise von einem Polizisten geschrieben wird, der den Hergang selbst nicht beobachtet hat. Stattdessen wird, um bei diesem Beispiel zu bleiben, der Polizist sich durch Befragung von Zeugen (Beobachtern) und Protagonisten nachträglich einen Überblick über den Ablauf verschaffen. Der Berichterstatter wird damit zum allwissenden Erzähler. ‘Allwissend’ bedeutet dabei nicht, daß der Berichterstatter letztendlich wirklich alle Hintergründe des Ablaufs kennt. Vielmehr soll mit diesem Begriff zum Ausdruck gebracht werden, daß der Berichterstatter zu Beginn des Erzählzeitraums den Verlauf des Ablaufs bereits kennt. Er ist dadurch nicht gezwungen, den Bericht chronologisch aufzubauen. Ein gutes Beispiel für diese Freiheit des Berichterstatters ergibt sich aus folgendem Satz der oben angegebenen ersten Beschreibung:

*It headed for the last filling place.*

Ein seine individuelle Wahrnehmung beschreibender Beobachter hätte an dieser Stelle der Beschreibung diese Aussage nicht treffen können, da das beschriebene Ziel, das Ansteuern eines bestimmten Tankplatzes, zum Erzählzeitpunkt noch gar nicht zu erkennen gewesen sein dürfte. Der ‘allwissende’ Berichterstatter ist, dadurch daß er den Verlauf des Ablaufs bereits kennt, jedoch in der Lage, in seiner Beschreibung Vorgriffe auf kommende Teilabläufe vorzunehmen. Da der Berichterstatter sich vor der Abfassung seines Berichts einen möglichst umfassenden Überblick über die Handlungsmotive seiner Hauptakteure verschafft haben dürfte, wird sein Bericht sich hauptsächlich auf diese Angaben stützen. Dies bedeutet, daß ein Bericht im Gegensatz zu szenischen Beschreibungen möglichst wenig Mutmaßungen und Wertungen des Erzählers enthalten, sondern sich durch ein gesteigertes Maß an ‘Sachlichkeit’ auszeichnen wird. Grundsätzlich ist der Berichterstatter, ähnlich wie der seine individuelle Wahrnehmung beschreibende Beobachter, frei in der Wahl seiner Hauptakteure. Allerdings dürfte diese Wahl durch den Zweck des Berichts und durch die Funktion des Berichterstatters beschränkt werden.

## 2.3 Perspektivität

Der vorige Abschnitt hat gezeigt, daß bei der natürlichsprachlichen Beschreibung von Realweltszenen drei Erzählstile zur Verfügung stehen. Dennoch ist es möglich, ein- und denselben Ablauf auf mehr als drei Arten zu beschreiben. Diese Vielfalt ergibt sich aus der *Perspektive*, mit der der Ablauf jeweils betrachtet wird.



### 2.3.1 Phänomenologische Herangehensweise

Werden die Begriffe ‘Standpunkt’ und ‘Perspektive’ genannt, wird der Leser diese Begriffe vermutlich zuerst räumlich interpretieren. Im folgenden ist ein Beispiel für eine natürlichsprachliche Äußerung gegeben, die eine räumliche Relation enthält.

(1) *Karlsruhe befindet sich **südlich von** Mannheim.*

Diese Äußerung ist immer wahr, unabhängig davon, wo sich der Urheber der Äußerung räumlich befindet. Im Gegensatz dazu muß, um den Wahrheitswert des folgenden Satzes bewerten zu können, der Standpunkt des Äußernden bekannt sein:

(2) *Karlsruhe befindet sich **hinter** Mannheim.*

Letztgenannter Satz impliziert also eine *Perspektive*. Der Begriff der *Perspektive* soll nun von dem Spezialfall der räumlichen Relationen auf allgemeinere natürlichsprachliche Äußerungen erweitert werden. Man betrachte dazu die folgenden beiden Beispielsätze:

(3) *Für die CDU ist das Wahlergebnis ein toller Erfolg.*

(4) *Das Wahlergebnis ist ein toller Erfolg.*

Analog zu (1) stellt auch (3) eine Äußerung dar, deren Wahrheitswert (weitestgehend) unabhängig von einer Perspektive bestimmt werden kann. Beispielsatz (4) hingegen enthält, wie auch Aussage (2), eine persönliche Perspektive des Äußernden. Es läßt sich feststellen, daß Äußerungen mit perspektivischem Charakter einen mehr oder minder bestimmten Standpunkt implizieren. Dies bedeutet zweierlei (vgl. [Canisius 87]):

1. Perspektivische Äußerungen sind nur aus der durch diesen Standpunkt bedingten Perspektive richtig und angemessen.
2. Perspektivische Äußerungen erlauben umgekehrt Rückschlüsse auf den persönlichen Standpunkt des Äußernden.

Im Gegensatz zu (2) ist in (4) der Begriff der *Perspektive* lediglich eine Metapher, da sich die Perspektive nicht räumlich, sondern lediglich ‘abstrakt’ bestimmen läßt. Die Einleitung von [Canisius 87] enthält einige interessante Ausführungen über den metaphorischen Charakter des Begriffs *Perspektive*, die allerdings aufgrund ihres Umfangs und ihrer Verästelung in andere Themengebiete an dieser Stelle nicht wiederholt werden. Zusammenfassend stellt der Autor fest, daß die Basis der Metaphorik von Standpunkt und Perspektive sich daraus zu ergeben scheint, daß der Mensch dazu neigt, erstens Abstraktes zu Gegenständen zu metaphorisieren, welche allgemein immer (mindestens) zwei Seiten haben. Zweitens metaphorisiert der Mensch persönliche

Einstellungen zu Sichtweisen, die dann, wie alles Sehen, drittens stets perspektivisch sind. Die Perspektive natürlichsprachlicher Äußerungen besitzt im übrigen nur dann Relevanz, wenn sich der Standpunkt des Urhebers der Äußerung von dem Standpunkt des Adressaten der Äußerung unterscheidet. Fragt beispielsweise ein Autofahrer seinen Beifahrer, ob Karlsruhe hinter Mannheim liege, so ist für beide Personen die Frage eindeutig zu beantworten. Die Perspektive spielt keine entscheidende Rolle, da die Perspektive beider Personen gleich ist. Relevant wird Perspektivität immer erst dann, wenn sich der Standpunkt des Äußernden von demjenigen des Empfängers der Äußerung unterscheidet.

In den bisherigen Ausführungen wurden Beispiele für unterschiedliche Perspektiven gezeigt, die sich aus räumlichen Relationen oder aus *Wertungen* (z.B. in (4) das Wahlergebnis als Erfolg oder Mißerfolg) ergeben. Es gibt noch weitere Ansatzpunkte für Perspektiven. So gibt es zum einen die *zeitliche Perspektive*, wie sie durch folgende Sätze verdeutlicht wird:

- (5) *The car was filled up with petrol. Then it left the pump.*
- (6) *The car left the pump after it had been filled up with petrol.*

Beide Sätze sagen inhaltlich das gleiche aus, unterscheiden sich jedoch in ihrem Referenzzeitpunkt.

Desweiteren bewirkt die Festlegung der Hauptakteure eines zu beschreibenden Ablaufs die Einnahme einer bestimmten Perspektive, wie folgende Beispielsätze zeigen (aus [Lindemann 87]):

- (7) *John hit Mary.*
- (8) *John hit his wife.*
- (9) *Mary's husband hit her.*

Wenn man davon ausgeht, daß in allen drei Sätzen vom gleichen Ehepaar John und Mary die Rede ist, so sagen alle drei Sätze inhaltlich das gleiche aus. Unterschiedlich ist jedoch jeweils die Perspektive. Steht in (8) mehr die Person 'John' im Vordergrund, so ist dies in (9) mehr die 'Mary', während in (7) beide Personen 'gleichberechtigter' erscheinen, also ein 'neutralerer' Standpunkt gewählt wurde.

Zusammenfassend läßt sich feststellen, daß sich der Begriff der *Perspektive* zumindest aus unterschiedlichen räumlichen (1,2), zeitlichen (5,6), wertenden (3,4) und empathischen (7-9) Standpunkten ergeben kann.

### 2.3.2 Logische Struktur der Perspektive

Haben wir uns bisher dem Begriff der Perspektive lediglich phänomenologisch genähert, so soll jetzt versucht werden, diesen Begriff seiner elementaren 'logischen' Struktur



Sammlung von beobachteten Tatsachen dar, die von einem allwissenden Erzähler vorgelegt wird. Im Idealfall sollten sich zwei Berichte, die zu demselben Ablauf erstellt wurden, lediglich im Wortlaut, nicht jedoch in ihrem Inhalt wesentlich unterscheiden. Daraus ergeben sich folgende Einschränkungen in der Wahl der Perspektive.

- Da in dem Bericht das Verhalten von Fahrzeugen, welche naturgemäß intrinsische Vorderseiten haben, in Straßenverkehrsszenen geschildert werden soll, wird grundsätzlich eine räumlich *intrinsische* Perspektive gewählt, und nur in Ausnahmefällen eine räumlich deiktische Sichtweise.
- Der Berichterstatter ist, wie bereits zuvor betont, grundsätzlich frei in der Wahl der zeitlichen Perspektive.
- Grundsätzlich ist der Berichterstatter in der Wahl seiner Hauptakteure, und damit in seiner empathischen Perspektive, frei.
- Da ein Bericht neutral und sachlich verfaßt werden soll, sind Wertungen die Ausnahme.

### 2.3.3.2 Szenische Beschreibung aus Beobachtersicht (SBB)

Szenische Beschreibungen schildern Abläufe inkrementell analog ihres zeitlichen Verlaufs und analog des Wissensstandes des Erzählers. Im Gegensatz zum Berichterstatter kann der erzählende Beobachter bei dieser Beschreibungsweise nicht auf Wissen zurückgreifen, das er erst im Nachhinein erschlossen haben könnte. Folglich kann er über Absichten der am Ablauf Beteiligten und über den weiteren Verlauf des Ablaufs nur Mutmaßungen anstellen. Dies bedeutet, daß szenische Erzählungen aus Beobachtersicht üblicherweise ein hohes Maß an subjektiver Wahrnehmung und Beurteilung durch den Beobachter enthalten. Dies ergibt folgende Beschränkungen in der Wahl der Perspektive.

- Analog zur Vorgehensweise beim Bericht wird grundsätzlich eine räumlich intrinsische Sichtweise gewählt, und nur in Ausnahmefällen eine deiktische Sichtweise.
- Die zeitliche Perspektive ist fest vorgegeben durch den Zwang zur streng chronologischen Schilderung des Ablaufs.
- Wie der Berichterstatter ist auch der erzählende Beobachter grundsätzlich frei in der Auswahl seiner Hauptakteure.
- Ein Beobachter beschreibt das Verhalten von Agenten zum Zeitpunkt der Beobachtung. Da er Ziele, Absichten und Motivation der Agenten nicht notwendigerweise kennt, ist er auf Mutmaßungen und eigene Schlußfolgerungen angewiesen. Diese können individuellen Wertmaßstäben unterliegen und sich somit von Beobachter zu Beobachter unterscheiden.

### 2.3.3.3 Szenische Beschreibung aus Protagonistensicht (SBP)

Bei einer szenischen Beschreibung aus Protagonistensicht schildert ein unmittelbar Beteiligter seine persönliche Wahrnehmung des Ablaufs. Er stellt die Motive seiner Handlung, seinen persönlichen Zustand, seine Absichten und seine Ziele dar. Desweiteren schildert er die Umstände, die zu Änderungen seines Zustands und seiner Handlungen führten. Der Protagonist weiß folglich mehr über die Hintergründe, die einen Ablauf beeinflussen, als ein unbeteiligter Beobachter. Allerdings kennt der Protagonist im Gegensatz zum 'allwissenden Erzähler' beim Bericht nur seine Sicht der Dinge, und nicht diejenige der anderen Beteiligten. Somit wird die Protagonistenperspektive folgendermaßen eingeschränkt.

- Zur Festlegung der räumlichen Perspektive wird analog zu den anderen Erzählstilen verfahren.
- Für die zeitliche Perspektive gelten dieselben Aussagen wie für szenische Beschreibungen aus Beobachtersicht.
- Im Gegensatz zum Berichterstatter und zum erzählenden Beobachter ist der Protagonist nicht frei in der Wahl seiner Hauptakteure und damit in seiner empathischen Perspektive. Einziger Hauptakteur ist hier der Protagonist selbst, der seine persönliche Sicht der Dinge schildert. Andere Agenten werden nur relativ zum eigenen Verhalten erwähnt.
- Ähnlich wie beim Beobachter ist die Einschätzung einer Szene durch einen Protagonisten abhängig von den individuellen Wertmaßstäben des Erzählers.

## 2.4 Erzählercharakteristiken

Wie bereits mehrfach betont, läßt sich ein und derselbe Ablauf auf vielerlei Arten schildern. Will man Abläufe auf eine bestimmte Art beschreiben, so sind vier Entscheidungen zu treffen.

1. Wahl des Erzählstils: Bericht, szenische Beschreibung aus Protagonistensicht oder szenische Beschreibung aus Beobachtersicht.
2. Festlegung der beschreibungsrelevanten Abläufe (z.B. im Diskursbereich Straßenverkehr: Überholvorgänge, Tankplatzbelegungen, mögliche Gefahrensituationen etc.).
3. Festlegung der Hauptakteure.
4. Bestimmung der Erzählercharakteristik.

Wie bereits erwähnt, ist zu beachten, daß bei der Wahl des Erzählstils ‘szenische Beschreibung aus Protagonistensicht’ der Protagonist als einziger Hauptakteur bereits implizit feststeht. Auch ist zu erwähnen, daß die Wahl der relevanten Abläufe und der Hauptakteure voneinander abhängen kann. Wird beispielsweise die Schilderung aller Überholvorgänge als Beschreibungsziel formuliert, so ergeben sich die Hauptakteure der Beschreibung unmittelbar: Hauptakteure sind all diejenigen Fahrzeuge, die in einen Überholvorgang verwickelt waren. Tabelle 2.2 gibt alle erlaubten Kombinationen der oben genannten Entscheidungsfaktoren an. Auf die Bedeutung der freien Erzählercharakteristiken soll an dieser Stelle noch genauer eingegangen werden. So wurden für szenische Beschreibungen in der Tabelle Wertmaßstäbe als freie Erzählercharakteristik angegeben. Dies bedeutet folgendes: Gegeben seien zwei verschiedene Erzähler, die vom selben Beobachtungspunkt aus einen bestimmten Ablauf mit bestimmten Hauptakteuren beschreiben sollen. Obwohl somit die Ausgangsbedingungen nach der zweiten, dritten und vierten Spalte der Tabelle identisch sind, ist zu erwarten, daß sich die Beschreibungen, die die beiden Erzähler erzeugen, nicht nur im Wortlaut voneinander unterscheiden. Diese Unterschiede resultieren aus verschiedenen Charakteristiken der Erzähler. Diese spiegeln sich beispielsweise in unterschiedlichen Wertungen derselben Abläufe wider, z.B. in Aussagen wie: *Das Verhalten des entgegenkommenden Fahrzeuges war verantwortungslos / mutig / völlig in Ordnung*. Diese Wertungen sind abhängig von der Persönlichkeit des erzählenden Protagonisten (Charakter, soziokulturelle Faktoren, Stimmungen etc.). Auch die *Funktion* und die *persönliche Bindung* des Erzählers fallen ins Gewicht. So bewertet ein und dieselbe Person den gleichen Ablauf vermutlich anders, je nachdem ob sie als Polizist oder als Sensationsreporter einen Unfall kommentiert (Einflußfaktor ‘Funktion’). Ebenso unterscheidet sich die Wertung desselben Ablaufs durch dieselbe Person, wenn der Beobachter z.B. bei einem Unfall lediglich Schaulustiger ist oder durch den Unfall eine geliebte Person verloren hat (Einflußfaktor ‘persönliche Bindung’).

## 2.5 Natürlichsprachliche Anfragen

Aus Tabelle 2.2 ergeben sich zehn Formen natürlichsprachlicher Beschreibungen. Vor Beginn der Beschreibungsgenerierung muß daher festgelegt werden, welche dieser zehn Varianten im Einzelfall vom Benutzer erwünscht ist. Wie bereits mehrfach ausgeführt, ist die natürliche Sprache für den Menschen ein sehr intuitives Kommunikationsmittel. Deshalb liegt es nahe, die gewünschte Beschreibungsvariante durch eine natürlichsprachliche Anfrage zu bestimmen. Tabelle 2.3 gibt für jede Beschreibungsvariante aus Tabelle 2.2 eine korrespondierende natürlichsprachliche Anfrage an. Diese Tabelle soll nicht implizieren, daß vom Benutzer verlangt wird, Anfragen exakt in den angegebenen Wortlauten zu formulieren. Wäre man als Konstrukteur so restriktiv, würde eine Menüauswahl mehr Sinn machen als vom Benutzer die Mühe abzuverlangen, seine Auswahl auch noch textuell exakt anzugeben. Tabelle 2.3 soll lediglich folgendes zeigen: Etabliert man ein solches Anfragesystem, so steht das System vor der Aufgabe,

Nr.	Erzählstil	Relevante Abläufe	Hauptakteure	Freie Erzählercharakteristik
1	Bericht	frei	frei	zeitlicher Aufbau
2		frei	gebunden	dto.
3		gebunden	frei	dto.
4		gebunden	gebunden	dto.
5	Szenische Beschreibung aus Protagonistensicht	frei	gebunden	Wertmaßstäbe
6		gebunden	gebunden	dto.
7	Szenische Beschreibung aus Beobachtersicht	frei	frei	Wertmaßstäbe
8		frei	gebunden	dto.
9		gebunden	frei	dto.
10		gebunden	gebunden	dto.

Tabelle 2.2: Auflistung aller erlaubten Konstellationen für natürlichsprachliche Beschreibungen. Die verschiedenen Konstellationen sind nummeriert. Die Nummern entsprechen denjenigen aus Tabelle 2.3. Nähere Erläuterungen im Text.

aus einer gegebenen Anfrage die ideale Beschreibungsvariante zu extrahieren. Dabei ist zu erwarten, daß Anfragen des Benutzers üblicherweise zumindest in dem Sinne unter-spezifiziert sein werden, daß der optimale Erzählstil des Antworttextes nicht explizit ableitbar sein wird. Dem Benutzer wird egal sein, wie der Erzählstil heißt, in dem die Antwort auf seine Anfrage formuliert wurde, solange die Antwort seine Anforderungen an Klarheit und Zuverlässigkeit erfüllt. Das System steht jedoch vor der Aufgabe, denjenigen Erzählstil zu finden, der eben jene Anforderungen am besten erfüllen kann. Generell sind für ein System zur Generierung natürlichsprachlicher Beschreibungen von Straßenverkehrsszenen durch Bildfolgenauswertung zwei Arbeitsszenarien denkbar.

- **Ablaufbeschreibung:** Bei der Ablaufbeschreibung stellt der Benutzer natürlichsprachliche Anfragen zur Auswertung einer gegebenen beobachteten Straßenverkehrsszene. Das System analysiert die Anfrage und leitet aus ihr die einzelnen

Nr.	Korrespondierende Anfrage
1	Was ist in der Szene alles passiert?
2	In welche Abläufe waren die von mir ausgewählten Agenten verwickelt?
3	Inwieweit waren welche Agenten in die von mir ausgewählten Abläufe verwickelt?
4	Inwieweit waren die von mir ausgewählten Agenten in die von mir ausgewählten Abläufe verwickelt?
5	Protagonist, wie hast du die Szene erlebt?
6	Protagonist, wie hast du die von mir ausgewählten Abläufe erlebt?
7	Beobachter, welche Eindrücke hast du im Verlauf der Szene gesammelt?
8	Wie lassen sich die Abläufe chronologisch schildern, in die die von mir ausgewählten Agenten verwickelt waren?
9	Wie sind die von mir ausgewählten Abläufe chronologisch verlaufen?
10	Wie sind die von mir ausgewählten Abläufe chronologisch für die von mir ausgewählten Agenten verlaufen?

Tabelle 2.3: Für jede Beschreibungsvariante aus Tabelle 2.2 ist eine korrespondierende Anfrage angegeben.

Parameter ab, die die Perspektive bestimmen, mit Hilfe derer die Szene ausgewertet werden soll.

- **Verkehrsszenenmassenauswertung:** Hierbei wird vom Benutzer eine Perspektive spezifiziert, mit Hilfe derer wechselnde Abläufe ausgewertet werden. Diese feststehende Perspektive ist folglich fester Bestandteil der Systemspezifikation und muß nicht aus jeder Anfrage neu berechnet werden.

Während folglich bei der Ablaufbeschreibung der Ablauf fest ist und die Perspektive wechselt, mit Hilfe derer der Ablauf betrachtet wird, steht bei der Verkehrsszenenmassenauswertung die Perspektive fest und die Abläufe wechseln. Beide Alternativen lassen sich in ein und demselben System verwirklichen.

- Bei der *Verkehrsszenenmassenauswertung* liegt ein fest definierter Erzähler vor. Die inhaltlichen Schwerpunkte, auf die dieser Erzähler besonders abhebt, sind durch die Spezifikation fest vorgegeben. Gleiches gilt für die Wertmaßstäbe, die der Erzähler zur Beurteilung von Abläufen anlegt. Gleiches gilt auch für den Erzähltyp, in dem die zu generierenden natürlichsprachlichen Beschreibungen formuliert werden. Ausgehend von der vorgegebenen Erzählerspezifikation werden



Straßenverkehrsszenen ausschließlich mit Hilfe der Perspektive ausgewertet, die sich aus der Erzählerspezifikation eindeutig ableiten läßt.

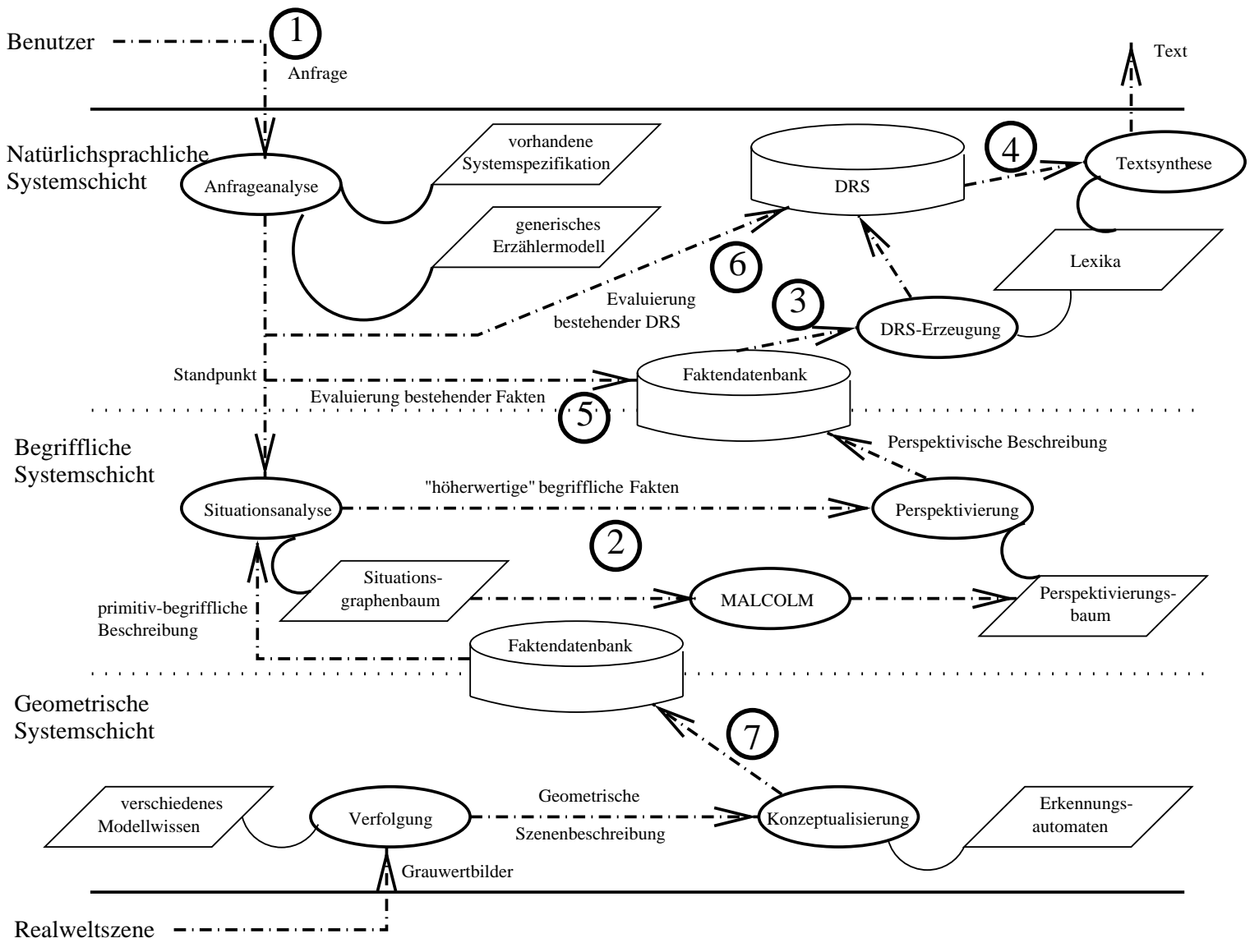
- Bei der *Ablaufbeschreibung* ist der Erzähler nicht fest spezifiziert, sondern definiert sich durch jede Anfrage neu. Bei der Analyse der Anfrage werden die Parameter bestimmt, die den Erzähler und damit dessen Perspektive beschreiben. Mit Hilfe dieser Perspektive wird ein gegebener Ablauf analysiert. Kann durch die Vagheit oder die Knappheit einer Anfrage die Perspektive nicht eindeutig festgelegt werden, so muß das System Mechanismen vorsehen, die die unvollständigen bzw. vagen Vorgaben zu einer konsistenten und pragmatischen Spezifikation vervollständigen (z.B. eigenständige Bestimmung des geeigneten Erzählstils durch das System). Ist die Perspektive ermittelt, so kann, analog zur Verkehrsszenenmassenauswertung beginnend bei der signalnahen Auswertung, prinzipiell die Berechnung der Beschreibung angegangen werden. Da jedoch bei der Ablaufbeschreibung mit immer neuen Anfragen bezüglich derselben Szene zu rechnen ist, erscheint es fraglich, ob es sinnvoll ist, die Berechnung bei jeder neuen Anfrage in der signalnahen Auswertung neu zu starten. Stattdessen erscheint es zumindest als zweckmäßig, einen Grundstock an Auswertungsergebnissen in einer Datenbank dauerhaft zu speichern, auf die im Bedarfsfalle bei weiteren Anfragen zugegriffen werden kann.

## 2.6 Systemaufbau

In den vorangegangenen Abschnitten wurde gezeigt, daß für jede Realweltszene mehr als eine als zutreffend zu erachtende natürlichsprachliche Beschreibung formulierbar ist. Die Gemeinsamkeiten und Unterschiede verschiedener natürlichsprachlicher Beschreibungen wurden herausgearbeitet. Ziel ist es nun, ein Bildfolgenauswertungssystem zu entwerfen, das im Prinzip alle Arten natürlichsprachlicher Beschreibungen extrahieren kann. Abbildung 2.2 zeigt einen solchen Systementwurf als Skizze.

Der Benutzer stellt eine (zum jetzigen Stand noch nicht natürlichsprachlich, sondern durch Angabe relevanter Parameter formulierte) Anfrage an das System. Die natürlichsprachliche Schicht analysiert diese Anfrage und bestimmt daraus die Parameter zur perspektivischen Inhaltssynthese (Ziffer '1' in Abbildung 2.2). Läßt die Anfrage auf eine Verkehrsszenenmassenauswertung schließen, so werden diese Parameter aus einer vorgegebenen Spezifikation abgeleitet. Andernfalls ergeben sich die Parameter aus der Anfrage selbst, wobei zur Charakterisierung des Erzählers auf ein Erzählermodell zugegriffen wird. Der ermittelte Parametersatz, der den Erzählerstandpunkt beschreibt, wird an die begriffliche Schicht weitergeleitet. Für jeden Hauptakteur wird eine *Situationsanalyse* durchgeführt. Bei der Perspektivierung werden aus der in der Situationsanalyse ermittelten Situationsfolge eine perspektivische Beschreibung berechnet (Ziffer '2'). Dazu wird ein Perspektivierungsbaum verwendet, der von dem Programm MALCOLM aus dem gegebenen Situationsgraphenbaum abgeleitet wird. Die perspektivische

Abbildung 2.2: Entwurf eines Systems zur Generierung natürlichsprachlicher Beschreibungen aus Bildfolgen. Nähere Erläuterungen im Text.



Beschreibung wird in die Faktendatenbank aufgenommen und an die natürlichsprachliche Schicht übergeben. Dort erfolgt deren Transformation in eine Diskursrepräsentationsstruktur (DRS, [Kamp&Reyle 93]), welche die Datenmenge unter semantischen Gesichtspunkten analysiert und interpretiert. In dieser Schicht sollen die stilistischen Entscheidungen zur Texterzeugung getroffen und die natürlichsprachliche Beschreibung erzeugt werden (4).

Zwei Umstände sind denkbar, die dieses Szenario im Einzelfall vereinfachen. Zum einen kann bei der Anfrageanalyse geprüft werden, ob die zur Beantwortung der Anfrage relevante Information bereits aufgrund früherer Anfragebearbeitungen Inhalt der semantischen Repräsentation in der natürlichsprachlichen Schicht ist (6). In diesem Falle kann die Antwort direkt aus dieser internen Repräsentation generiert werden, ohne daß auf die begriffliche Auswertungsstufe zurückgegriffen werden müßte. Zum anderen ist denkbar, daß die zur Beantwortung der Anfrage relevante Information zwar nicht Bestandteil der internen Repräsentation in der natürlichsprachlichen Schicht, jedoch bereits Bestandteil der Faktendatenbank der begrifflichen Schicht ist (5). In diesem Falle ist eine erneute Situationsanalyse überflüssig, weil die Antwort bereits generiert werden kann, indem die schon vorhandenen Fakten in der Faktendatenbank an die natürlichsprachliche Schicht weitergereicht werden.

Die Situationsanalyse in der begrifflichen Schicht erfolgt anhand einer primitiv-begrifflichen Beschreibung, die bereits a-priori aus der geometrischen Szenenbeschreibung der signalnahen Auswertungsschicht abgeleitet wird (7). Natürlich ist es denkbar, diese primitiv-begriffliche Beschreibung nicht komplett, sondern nur in dem Umfang zu berechnen, wie es eine gegebene Anfrage verlangt. Ebenso ist denkbar, daß jede einzelne Anfrage die signalnahe Auswertung anstößt, so daß Fahrzeuge nur im Bedarfsfalle verfolgt werden. Dieser Ansatz ist jedoch nicht Gegenstand der vorliegenden Arbeit und des in Abbildung 2.2 gezeigten Systementwurfs. Stattdessen wird aus Gründen der vorläufigen Problemvereinfachung bei der Anfragebearbeitung auf eine bereits vollständig vorhandene, aber nichtsdestoweniger auf Realdaten basierende primitiv-begriffliche Beschreibung zugegriffen.

## 2.7 Geometrische Systemschicht

Die begriffliche und die natürlichsprachliche Schicht (siehe Abbildung 2.2) werden in den folgenden beiden Kapiteln realisiert. In dem vorliegenden Abschnitt wird die geometrische Systemschicht vorgestellt, die Ausgangspunkt der vorliegenden Arbeit ist ([Koller et al. 93], [Kollnig&Nagel 97], [Haag 98], [Haag&Nagel 98a], [Haag&Nagel 99]). [Nagel 77] enthält erste Ansätze zur Implementierung eines Systems zur Auswertung von Bildfolgen, die mit einer stationären Schwarz/Weiß-Kamera aufgenommen wurden. [Nagel 81] liefert einen Überblick über Anwendungsgebiete, grundlegende Probleme und Aufgaben der signalnahen Bildauswertung sowie die Vorstellung und den Vergleich unterschiedlicher Lösungsansätze und Auswertungstechniken. Durch Rasterung

der Bildebene und Quantisierung der Grauwerte eines zweidimensionalen Grauwertverlaufs läßt sich ein digitales Bild erzeugen. Durch Diskretisierung der Zeitachse derart, daß zu jedem Punkt des Zeitrasters ('Zeitpunkt') ein digitales Bild erfaßt wird, läßt sich die zeitliche Änderung von Grauwerten erfassen. Auf diese Weise entsteht eine zeitliche Bildfolge.

Die geometrische Schicht berechnet eine geometrische Zustandsschätzung für alle relevanten Objekte der Szene und für jeden Halbbildzeitpunkt. Dabei werden statische Szenenkomponenten wie Fahrbahnen und Straßenverkehrsschilder explizit modelliert. Die geometrische Beschreibung des Verhaltens von Fahrzeugen wird automatisch geschätzt. Die Verfolgung sich bewegender Objekte wird automatisch initialisiert durch die Segmentierung optischer Flußfelder und durch die Einbeziehung von Wissen über die Kameraposition sowie die Straßenführung. Das Wissen um die Kamera wird benutzt, um die segmentierten und in einer Ballung zusammengefaßten optischen Flußvektoren (*Objektbildkandidaten* genannt) in den Szenenbereich zu projizieren. Bei der Transformation vom Bild- in den Szenenbereich werden zu den Objektbildkandidaten (zur Zeit immer noch interaktiv) Straßenverkehrsfahrzeuge bestimmten Typs assoziiert, welche als *Objektkandidaten* verfolgt werden. Der Verfolgungsprozeß selbst besteht aus zwei Unterprozessen, welche beide auf Kalman-Filterung beruhen: Ein *Aktualisierungsschritt* basierend auf einem aktuellen Bild und ein *Prädiktionsschritt* für das nachfolgende Bild. Für den Aktualisierungsschritt, der den Zustandsvektor des Objektkandidaten iterativ an die Bildbereichshinweise angleicht, werden verschiedene Techniken benutzt, die den Gebrauch von Kantenelementen und Optischen-Fluß-Vektoren kombinieren. Für den Prädiktionsschritt wird ein Bewegungsmodell verwendet, womit die zu erwartende Position des Fahrzeugs im nächsten Halbbild bestimmt wird. Die geometrische Zustandsschätzung besteht aus den geschätzten (x,y)-Positionen des Fahrzeugs auf der Fahrbahnebene der Szene, aus seiner Orientierung, seiner Geschwindigkeit und seiner Winkelgeschwindigkeit. Zu diesen Trajektoriendaten werden unscharfe Attribute assoziiert. Zum Beispiel kann man von der geschätzten Geschwindigkeit durch die Assoziierung von Begriffen wie *langsam*, *normal* oder *schnell* abstrahieren, wobei ein Zusicherungsgrad angegeben wird, der aussagt, wie gut ein bestimmter Begriff zu der geometrischen Schätzung paßt. In ähnlicher Weise kann man begrifflich von der geschätzten Orientierung abstrahieren und die geschätzte Fahrzeugposition zur Assoziierung räumlicher Relationen bezüglich des Fahrbahnmodells verwenden. Da die begrifflichen Beschreibungen, wie bereits ausgeführt, einen Zusicherungsgrad und ein Gültigkeitszeitintervall beinhalten, können Beschreibungen in dieser Notation als Menge unscharfer metrisch-temporaler Logikfakten aufgefaßt werden.

[Heinze et al. 91] liefert einen ersten Ansatz zur Berechnung begrifflicher Attribute aus Trajektoriendaten. [Kollnig&Nagel 93] berechnen begriffliche Beschreibungen für in Bildfolgen beobachtete Geschehen. Dabei beschreibt ein Geschehen einen elementaren Ablauf im Sekundenbereich. Geschehen werden mit Bewegungsverben oder Verbalphrasen beschrieben. [Damm 93] überträgt die Geschehensdefinition, die auf deutschsprachigen Bewegungsverben und Verbalphrasen basiert, in eine englischsprachige Begriffs-

welt. Geschehen lassen sich im Hinblick auf mögliche Bezugsobjekte in *Agensgeschehen*, *Fahrbahn-*, *Objekt-* und *Ortsgeschehen* unterteilen. [Gerber&Nagel 96a] unterteilen Agensgeschehen in weitere Unterklassen und verknüpfen diese Klassifikation mit einem heuristischen Verfahren zur Ableitung komplexerer Beschreibungen. [Damm 94] erweitert die Geschehensdefinition um typische Geschehen, die sich an einer Tankstelle beobachten lassen (*Tankstellengeschehen*). [Mück 94] entwirft zum einen eine geometrische Beschreibung von innerstädtischen Straßenkreuzungen, zum anderen trägt er aus entsprechender Fachliteratur die verkehrswissenschaftlichen Begriffe zusammen, die zum Aufbau einer begrifflichen Beschreibung solcher Kreuzungen erforderlich sind. Die Zusammenhänge zwischen diesen einzelnen Begriffen stellt dieser Autor anhand von Begriffsgraphen durch Aggregations- und Generalisierungsbeziehungen dar. Dabei korrespondieren das geometrische und das begriffliche Modell in wichtigen Komponenten. [Fautz 96] verknüpft die geometrische und die begriffliche Beschreibung, indem er im ersten Schritt Fahrbahnobjekte anhand von Bildbereichsmerkmalen erkennt. Diese Fahrbahnobjekte werden als *atomar*, d.h. nicht weiter untergliederbar betrachtet, und durch Rückgriff auf das begriffliche Modell zu komplexeren Strukturen aggregiert. Damit ist die Voraussetzung geschaffen, um Kenntnis über die Semantik bestimmter Fahrbahnanteile in der Szene zu erlangen, und diese Kenntnis für die Beschreibung des Verhaltens eines Fahrzeuges, das sich auf besagtem Fahrbahnanteil bewegt, zu nutzen. [Damm 94] erstellt, basierend auf [Nagel 91], ein begriffliches Fahrbahnmodell für eine Tankstelle. Die begriffliche Beschreibung der Fahrbahn stimmt in ihrer Notation mit der begrifflichen Beschreibung von Fahrzeugverhalten überein und kann deshalb ebenfalls als Menge unscharf metrisch-temporaler Logikfakten aufgefaßt werden.

Zur erfolgreichen Verfolgung von Fahrzeugen auch unter erschwerten Auswertungsbedingungen, wurde das System u.a. um Ansätze erweitert, die die explizite Modellierung von Verdeckungen erlauben ([Frank 95], [Frank et al. 96], [Haag et al. 97]). Die Autoren erweitern die begriffliche Schicht um eine Terminologie zur Beschreibung von Verdeckungen. [Pernozzoli 96] nutzt diese zur Ableitung entsprechender natürlichsprachlicher Beschreibungen. [Schwarz 97] erweitert den Verfolgungsansatz auf mehrgliedrige Fahrzeuge. [Stockmar 99] untersucht zur Effizienzsteigerung Möglichkeiten zur Parallelverarbeitung bei der Bildauswertung.

# Kapitel 3

## Begriffliche Abstraktion

Die begriffliche Schicht verwaltet die Faktendatenbank und abstrahiert, im Rahmen dieser Arbeit zum Zwecke der Generierung einer aussagekräftigen und gut strukturierten natürlichsprachlichen Beschreibung, begriffliche Primitiven zu komplexeren begrifflichen Beschreibungen. Sämtliche Abstraktionsschritte werden durch unscharfe, metrisch-temporale Logikprogramme realisiert (FMTL, [Schäfer 96]). Die von diesem Autor entwickelte Logiksprache *F-Limette* wird im nächsten Abschnitt vorgestellt. Darauf folgend wird erläutert, wie diese Sprache zur Abstraktion der auf der geometrischen Ebene extrahierten primitiv-begrifflichen Zustandsschätzung eingesetzt wird.

### 3.1 Literaturüberblick

[Neumann&Novak 86] verwenden in ihrem System NAOS eine begriffliche Hierarchie von Ereignissen und repräsentieren Wissen über relevante Veränderungen in einer Szene durch generische Ereignismodelle. Damit sollen in der (synthetisch erzeugten) geometrischen Szenenbeschreibung vorliegende Daten mit qualitativen, begrifflichen Beschreibungsformen assoziiert werden. Diese Vorgehensweise ist wie folgt zu kritisieren (vgl. [Nagel 88]). Die Arbeitshypothese, daß Ereignisse hierarchisch definiert werden können, ist praktisch nicht umsetzbar. Beispielsweise läßt sich ‘*Rennen*’ nicht als ‘*Gehen im oberen Geschwindigkeitsbereich*’ beschreiben, da das Gehen eines Sportlers durchaus schneller sein kann als das Rennen eines alten Mannes. Zur Abgrenzung von *Gehen* und *Rennen* ist deshalb eine Gangartenerkennung unter Rückgriff auf entsprechendes Modellwissen erforderlich. Dies zeigt, daß zur Erzeugung natürlichsprachlicher Beschreibungen Wissen über die jeweils einsetzbare Terminologie vorhanden sein muß. Dieses Wissen einzubringen, erscheint zumindest schwierig, wenn unmittelbar aus einer rein geometrischen eine natürlichsprachliche Szenenbeschreibung abgeleitet werden soll. Zum anderen fällt negativ in Betracht, daß die Systemleistung von NAOS aufgrund der Fülle erforderlicher Bildauswertungsdaten nur an synthetischen Daten getestet werden konnte, so daß viele Probleme der ‘realen Welt’ dort keine angemessene Berücksichti-

$\delta$	weak	medium	strong
$\leftarrow$	$\min(1, 1 - \wp_2 + \wp_1)$	$1 - \wp_2 + \wp_1 * \wp_2$	$\max(\wp_1, 1 - \wp_2)$
$\wedge$	$\min(\wp_1, \wp_2)$	$\wp_1 * \wp_2$	$\max(0, \wp_1 + \wp_2 - 1)$
$\vee$	$\min(1, \wp_1 + \wp_2)$	$\wp_1 + \wp_2 - \wp_1 * \wp_2$	$\max(\wp_1, \wp_2)$

Tabelle 3.1: Resultierender Wahrheitswert einer Unscharfen-Logik-Operation  $\delta$ , welche zwei Formeln  $\mathcal{F}_1, \mathcal{F}_2$  mit zugehörigen Wahrheitswerten  $\wp_1, \wp_2$  verknüpft, unterschieden nach schwacher (w), mittlerer (m) und starker (s) Semantik.

gung finden können.

Die in [Huang et al. 94], [Howarth&Buxton 98] und [Intille 98] beschriebenen Systeme besitzen eine Auswertungsstufe, die mit der begrifflichen Schicht der vorliegenden Arbeit vergleichbar ist. Alle drei verwenden Bayes'sche Wahrscheinlichkeitsnetzwerke ([Pearl 88], [Charniak 91]), um Realwelt-Straßenverkehrsszenen bzw. Fußballspiel-szenen zu interpretieren. In [Huang et al. 94] wird zu jedem erkannten Fahrzeug ein Wahrscheinlichkeitsnetzwerk assoziiert, welches Knoten enthält, die zu bestimmten Sensordaten korrespondieren, und Knoten, welche zu Beschreibungen korrespondieren, die aus diesen Daten abgeleitet wurden. Das Wissen über den Diskursbereich geht als Wahrscheinlichkeitswerte, welche mit jedem Knoten assoziiert werden, in das Netzwerk ein. Diese Wahrscheinlichkeitswerte können gelernt werden, sofern ein ausreichend großer Trainingsdatensatz vorhanden ist. Die Verwendung Bayes'scher Netzwerke fördert nicht die intuitive Repräsentation von Wissen. Die Inspektion und Erweiterung eines solchen Netzwerkes ist besonders im Fehlerfalle aufgrund der verwobenen stochastischen Abhängigkeiten sehr schwierig. Die Performanz des Netzwerkes ist zum einen von der Wahl der Wahrscheinlichkeitswerte abhängig, für deren gezielte Festlegung in der Regel keine fundierte Theorie vorhanden ist. Die Effizienz der Ausprägung eines Bayes'schen Netzwerkes hängt andererseits stark von der Struktur desselben ab, so daß die Netzwerke oft nicht ausschließlich nach Vorgaben strukturiert werden, die sich aus der 'realen Welt' ergeben, sondern 'diskursbereichsfremde' Effizienzüberlegungen bei der Strukturierung Einfluß nehmen. Die Zweckmäßigkeit bayes'scher Verfahren zur Lösung praktischer Probleme ist in der Fachwelt umstritten ([Hearst 97]).

Diesen Nachteilen wird in der vorliegenden Arbeit mit der Bereitstellung eines begrifflichen, hierarchischen, inkrementellen und logikbasierten Formalismus zu begegnen versucht.

## 3.2 Unscharf metrisch temporale Logik

Logik genießt als Mittel zur Wissensrepräsentation und Inferenz weitverbreitete Akzeptanz. Viele gängige Repräsentationsformalismen lassen sich beispielsweise auf die

	LIMETTE-Ausdruck	korrespondierende FMTL-Formel
Faktum	$\lambda \mid t_1 : t_2 ! Rel.$	$\downarrow_\lambda \square_{[t_1, t_2]} Rel$
Regel	$\lambda \mid t_1 : t_2 ! (Rel_1 :- Rel_2).$	$\downarrow_\lambda \square_{[t_1, t_2]} (Rel_1 \leftarrow Rel_2)$
Anfrage	$\lambda \mid t_1 : t_2 ? Rel.$	$\downarrow_\lambda \diamond_{[t_1, t_2]} Rel$

Tabelle 3.2: Allgemeiner Aufbau von Fakten, Regeln und Anfragen in F-Limette und dazu korrespondierende FMTL-Formel. Zusicherungsgrade werden mit  $\lambda$  bezeichnet,  $t_1$  und  $t_2$  bezeichnen Zeitpunkte und  $Rel$  Relationen.

Prädikatenlogik erster Stufe (PL1) zurückführen. [Schäfer 96] erweitert die PL1 schrittweise um eine zeitliche und eine unscharfe Komponente. Zeitliches Schlußfolgern ist im Rahmen der vorliegenden Arbeit deswegen erforderlich, weil durch Bildfolgenauswertung gewonnene Aussagen in der Regel nur für bestimmte Zeitintervalle oder gar einzelne Zeitpunkte (Halbbildnummern) gültig sind. Da zum einen die Signalerfassung stochastischen Einflüssen unterliegt, die Assoziation begrifflicher Beschreibungen zu geometrischen Daten zum anderen unscharf erfolgt und natürlichsprachliche Begriffe generell einer gewissen Vagheit unterliegen, muß der Abstraktionsformalismus ebenso unscharfes Schlußfolgern auf der Basis nicht völlig unwahrer oder wahrer begrifflicher Aussagen vorsehen. Dies wird durch die FMTL ebenfalls gewährleistet.

*Temporale Logik* generalisiert die Prädikatenlogik durch die Hinzunahme von Zeitpunktmengen und der Definition einer partiellen Ordnung bezüglich dieser Mengen, wodurch die Wahrheitswerte einer Formel im Verlauf der Zeit variieren können. *Metrisch* temporale Logik basiert auf einer linearen, diskreten Zeitstruktur, nämlich derjenigen der ganzen Zahlen. In unserer Arbeit entspricht die betrachtete Zeitpunktmenge der Menge der Halbbildnummern der Videobilddaten. [Schäfer 96] nutzt die metrisch temporale Logik MTL ([Broszka 94]), in welcher zwei temporale Logikoperatoren ( $\square_{a,e}$  bzw.  $\diamond_{a,e}$ ) zur Denotation der universellen bzw. existentiellen Gültigkeit von Formeln innerhalb betrachteter Zeitpunktmengen eingeführt werden. Um die effiziente Behandlung von Regeln und Anfragen der MTL zu ermöglichen, schränkt [Schäfer 96] sie auf ihr hornlogisches Fragment *MTHL* ein. Er erweitert den Tableauekalkül *TAB* für das hornlogische Fragment der PL1 zu einem Tableauekalkül der *MTHL* und weist die Korrektheit und Vollständigkeit dieses Kalküls nach.

*Unscharfe Logik* generalisiert die Prädikatenlogik durch die Zulassung von reellwertigen Wahrheitswerten im geschlossenen Intervall  $[0,1] \subset \mathbb{R}$ . Demzufolge werden durch die vier Logikoperationen Subjunktion, Konjunktion, Disjunktion und Negation ebenfalls reellwertige Wahrheitswerte ermittelt. Tabelle 3.1 zeigt jeweils drei mögliche Semantiken (*schwach* ( $w$ ), *mittel* ( $m$ ) und *stark* ( $s$ )) bezüglich der drei erstgenannten Logikoperatoren. Die beiden einstelligen Logikoperatoren  $\downarrow_\lambda$  bzw.  $\uparrow_\lambda$  repräsentieren *Abschwächung* bzw. *Intensivierung* eines Zusicherungsgrads. [Schäfer 96] nutzt die unscharfe Logik *FL1*, deren Ausdrücke er wiederum auf ihr hornlogisches Fragment *FHL* einschränkt. Er erweitert *TAB* nun auch zu einem Tableauekalkül der *FHL* und weist



abermals die Korrektheit und die Vollständigkeit des so entstandenen Kalküls nach.

Im nächsten Schritt vereinigt dieser Autor die Eigenschaften der MTHL und der FHL zur unscharf metrisch temporalen Hornlogik *FMTHL*. Seine Logikprogrammiersprache *F-Limette* operationalisiert den Kalkül TAB der FMTHL. Somit ist keine Rücktransformation der FMTHL-Programms nach *PROLOG* erforderlich, wodurch die dynamische Veränderung eines Logikprogramms während seiner Ausführung ermöglicht wird.

Die in der begrifflichen Schicht vorgenommenen Berechnungen basieren auf der FMTHL. Logikprogramme bestehen im allgemeinen aus *Regeln*, anhand derer zugehörige *Anfragen* bearbeitet werden. Eine Regel gibt für ein Prädikat  $p$  (den *Kopf* der Regel) eine Bedingung an, aus deren Gültigkeit oder Ungültigkeit die Gültigkeit bzw. Ungültigkeit von  $p$  abzuleiten ist. Eine solche - als Logikformel formulierte - Bedingung wird als *Rumpf* der Regel bezeichnet. Ein Prädikat  $p$  kann zugleich Kopf mehrerer Regeln sein. Liegt eine Anfrage an das Logiksystem vor, so versucht dieses, anhand der vorgegebenen Regeln eine *Substitution* (Variablenbelegung) zu finden, durch welche sich die Gültigkeit der Anfrage ergibt. Dabei werden rekursiv für jedes abzuleitende Prädikat der Anfrage die Rümpfe aller Regeln, die dieses Prädikat als Kopf enthalten, auf ihre Ableitbarkeit hin überprüft. Tabelle 3.2 gibt den allgemeinen Aufbau von Ausdrücken in F-Limette und dazu korrespondierende FMTL-Formeln an. Fakten und Regeln werden generell universell quantifiziert, während Anfragen existentiell quantifiziert werden. Wird in einem F-Limette-Ausdruck der Zusicherungsgrad  $\lambda$  weggelassen, so wird stets absolute Gültigkeit vorausgesetzt. Für zeitlich unbeschränkt gültige Ausdrücke werden als Gültigkeitsgrenzen  $-i$  (*Infinity*) bzw.  $+i$  verwendet.

Mit Hilfe der FMTHL wird das Wissen aus verschiedenen Wissensquellen mit den Bildfolgenauswertungsergebnissen als begriffliche Zustandsschätzung in Form von Regeln und Fakten vereint. F-Limette ermöglicht die Wahl zwischen mehreren Inferenzstrategien. Die Suchstrategie *Tiefensuche* wählt unter mehreren Alternativen eine aus und verfolgt diese so lange, bis sie zum Ziel führt oder scheitert. Erst dann wird eine andere Alternative behandelt. Bei der *Breitensuche* werden mehrere Alternativen im Zeitmultiplex verfolgt, indem diese abwechselnd schrittweise vorgebracht werden. Diese Strategie ist aufwendiger als die Tiefensuche, gewährleistet aber, daß keine Pfade, die zu einer erfolgreichen Suche führen würden, unbearbeitet gelassen werden. Die *Bündelsuche* realisiert eine Breitensuche, wobei der Suchraum durch Heuristiken eingeschränkt wird. Überschreitet die Anzahl der zu verfolgenden Pfade einen gewissen Schwellwert, so wird die Zahl der weiterhin zu untersuchenden Pfade reduziert. Dabei wird jedem Pfad ein Gütewert zugeordnet, wobei Pfade mit hoher Güte bevorzugt werden. [Schäfer 96] gibt vier Heuristiken für die Pfadentscheidung bei einer Bündelsuche an und erlaubt zudem die Einführung weiterer, benutzerdefinierter, Heuristiken.

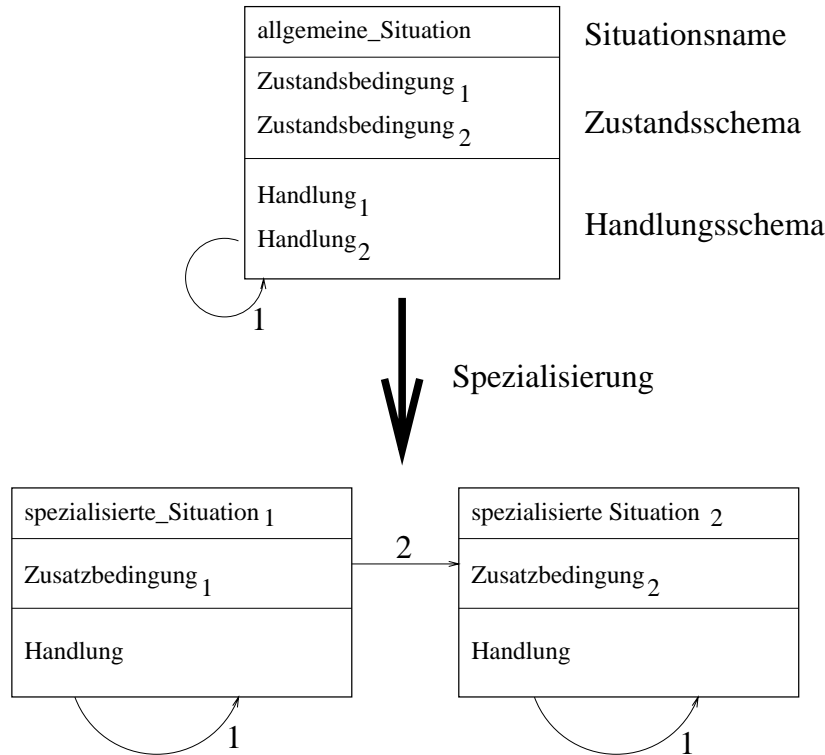


Abbildung 3.1: Allgemeiner Aufbau eines Situationsgraphenbaums. Die Knoten enthalten Situationsbeschreibungen, welche aus einem Situationsnamen und einem Zustandsschema bestehen. Die dünnen gerichteten Kanten geben Nachfolgebeziehungen zwischen Situationen derselben Hierarchiestufe an, wobei die neben den Kanten stehenden Zahlen deren Prioritäten bezeichnen ('1' entspricht hoher Priorität). Dicke gerichtete Kanten markieren Spezialisierungen zwischen Situationen. Nähere Erläuterungen im Text.

### 3.3 Situationsanalyse

Die *Situationsanalyse* beschäftigt sich mit der Bestimmung von Situationen handlungsfähiger Individuen (*Agenten*) in einer Szene abhängig vom eigenen Zustand des Individuums und vom Zustand seiner Umgebung. [Nagel 88] führt den Begriff *Situation* als begriffliche Einheit zur Wissensrepräsentation ein, durch welchen nicht nur Bewegungszustände, sondern auch zielgerichtete Aspekte (*Absichten*) beschrieben werden können. *Generisch beschriebene Situationen* repräsentieren zeitliche, räumliche und zielgerichtete Aspekte einer Szene.

[Krüger 91] verwendet *Situationsgraphenbäume* zur Darstellung von Situationen. Abbildung 3.1 zeigt den allgemeinen Aufbau eines Situationsgraphenbaums.

- Eine *Situationsbeschreibung* besteht aus mehreren Komponenten:
  - Die Situation wird anhand ihres *Situationsnamens* benannt.

- Das *Zustandsschema* beschreibt in generischer Form einen Zustand, den es im Rahmen eines Interpretationsprozesses auszuprägen gilt.
  - Das *Handlungsschema* enthält mit der Situation assoziierte Handlungen des Agenten oder Handlungserwartungen. Ist eine Situation vollständig spezialisiert, werden die in der zugehörigen Situationsbeschreibung definierten Handlungen ausgeführt (nichtinkrementelle Handlungsausführung). Zusätzlich kann verfügt werden, daß auch Handlungen hierarchisch übergeordneter Situationen (siehe unten) ausgeführt werden (inkrementelle Handlungsausführung).
  - Eine Situationsbeschreibung enthält zudem eine nach Priorität geordnete *Liste möglicher Nachfolgesituationen* der Situation.
- Ein *Situationsgraph* enthält als Knoten Situationsbeschreibungen. Die Liste der möglichen Nachfolgesituationen bildet die gerichteten Kanten dieses Graphen und benennt mögliche Prädiktionsbeziehungen von einem Zeitpunkt  $t$  zu einer Nachfolgesituation zum Zeitpunkt  $t + 1$ .
  - Situationen können zu untergeordneten Situationen spezialisiert werden. Gerichtete Spezialisierungskanten verbinden eine Situation höherer Hierarchiestufe mit einem untergeordneten Situationsgraphen. Auf diese Weise entsteht ein hierarchischer *Situationsgraphenbaum*. Speziellere Situationen erben die generischen Zustands- und Handlungsschemata ihrer übergeordneten Situationen.

[Krüger 91] beschreibt Zustands- und Handlungsschemata anhand von Begriffsgraphen ([Sowa 84]). Dabei handelt es sich um bipartite Graphen mit Individuen- und Relationsknoten. Relationsknoten spezifizieren Relationen zwischen den in den Individuenknoten repräsentierten Individuen. Begriffsgraphen lassen sich in prädikatenlogische Ausdrücke überführen. Der Vorteil des Begriffsgraphenformalismus liegt vor allem in der Möglichkeit zur grafischen Veranschaulichung komplexer relationaler Zusammenhänge. Das Programm SITRA ([Krüger 91]) stellt die Implementierung des vorangehend beschriebenen Ansatzes dar. [Schäfer 96] erkennt einige Nachteile, die sich aus dem prozeduralen Ansatz von [Krüger 91] ergeben, im Einzelnen:

- Die von [Krüger 91] gewählten Relationen beruhen allein auf einer zweiwertigen Logik und berücksichtigen somit keine unscharfen Beschreibungen.
- Eine einmal erfolgte Generalisierung einer Situation ermöglicht nicht die *bevorzugte* Rückkehr zur spezielleren Situation bei deren wiedererlangten Ausprägbarkeit.
- Durch die nichtrevidierbare Auswahl einer Situation kann die Traversierung leicht fehlgeleitet werden.
- Durch die - im allgemeinen starke - Überabtastung der Zustandserfassung bei der Situationsanalyse in der Bildfolgenauswertung werden zudem die ausprägbaren

Pfade länger und die Situationsanalyse selbst ineffizient. Diese Überabtastung ergibt sich dadurch, daß die Zeitskala der Beobachtungen mit den Bildnummern der zeitdiskreten Bildfolge gleichgesetzt wird. Als Konsequenz müssen Situationen des Situationsgraphen als mögliche Nachfolgesituation zu sich selbst modelliert werden. Jede Situation kann dann beliebig lange ausgeprägt werden. Umgekehrt erfordert die Ausprägbarkeit einer Nachfolgesituation bereits für einen einzigen isolierten Zeitpunkt die Berücksichtigung dieser Prädiktion.

[Zeier 99] vergleicht verschiedene Ansätze zur Situationsbeschreibung und entwirft eigene Situationsschemata zur Beschreibung von Situationsfolgen. [Schäfer 96] überträgt den prozeduralen Ansatz von [Krüger 91] auf einen logikbasierten Ansatz. Er beschreibt Situationsgraphenbäume fortan mit der von ihm entwickelten Situationsbeschreibungssprache *SIT++*. In der Sprache *SIT++* kodierte Situationsgraphenbäume lassen sich unmittelbar in Programme der Sprache *F-Limette* übersetzen. Das Zustandsschema einer Situation wird fortan mit konjunktiven Verknüpfungen unscharf metrisch temporaler Logikprädikate beschrieben, welche zur erfolgreichen Assoziation eines betrachteten Individuums zu einer Situation anhand entsprechender Fakten ausgeprägt werden müssen. Die Traversierung eines Situationsgraphenbaumes wird durch eine logische Anfrage gestartet. Die Traversierung selbst ist somit äquivalent zur Suche nach erfolgreichen Ableitungen einer entsprechenden Anfrage durch ein Logikprogramm. Die Suche erfolgt anhand der aus der signalnahen Schicht berechneten primitiven begrifflichen Szenenbeschreibung. Zudem erweitert [Schäfer 96] den Formalismus von [Krüger 91] um besondere unscharfe und zeitliche Aspekte. So können für die Situation selbst wie auch für die einzelnen Prädikate ihres Zustandsschemas Mindestzusicherungsgrade angegeben werden, deren Einhaltung bei der Ausprägung gewährleistet sein muß. Zudem erweitert [Schäfer 96] den Situationsgraphenformalismus um die Angabe von Mindest- und Höchstdauern für die aufeinanderfolgende Ausprägung einzelner Situationsbeschreibungen. Außerdem führt er Mindest- und Maximalverweilzeiten ein, für die eine zeitliche Prädiktion Bestand haben muß. Durch die Verwendung negativer Verweilzeiten können sogar zeitliche Überlappungen aufeinanderfolgender Situationen erreicht werden. Die Baumtraversierung kann mit unterschiedlichen Traversierungsstrategien durchgeführt werden. Zu Beginn der Traversierung wird die speziellste ausprägbare Situation gesucht. Im nächsten Zeitpunkt wird ausgehend von dieser Situation anhand der Prioritäten ihrer Nachfolgekanten eine Nachfolgesituation auf gleicher Hierarchieebene gesucht und ggf. noch weiter spezialisiert. Läßt sich eine solche Nachfolgesituation jedoch nicht finden, so wird die Auswertung in der nächstallgemeineren Situation fortgesetzt und dort nach einer ausprägbaren Nachfolgesituation gesucht. Die Konstruktion von Situationsgraphen zur Repräsentation üblicher Abläufe in Straßenverkehrsszenen ist ein Aspekt der vorliegenden Arbeit.

## 3.4 Perspektivische Situationsbeschreibung

In diesem Abschnitt wird beschrieben, wie aus einer primitiv-begrifflichen Beschreibung eine perspektivische begriffliche Beschreibung abgeleitet wird. Die Auswertung erfolgt dabei in drei Stufen. Im ersten Schritt wird aus Benutzer- oder Spezifikationsvorgaben der Erzählerstandpunkt bestimmt. Dann wird, noch unabhängig von der sich aus dem Erzählerstandpunkt ergebenden Perspektive, für jeden Hauptakteur der Beschreibung und für jeden Halbbildzeitpunkt die aktuelle Situation des Hauptakteurs bestimmt. Aus der so ermittelten Folge von Situationen werden, basierend auf dem ermittelten Erzählerstandpunkt, perspektivische begriffliche Beschreibungen abgeleitet. Die soeben angegebenen drei Schritte werden im folgenden detailliert ausgeführt.

### 3.4.1 Bestimmung des Erzählerstandpunkts

Der Erzählerstandpunkt  $\mathcal{S} = \langle H, E \rangle$  ergibt sich aus Benutzer- oder Spezifikationsvorgaben. Dabei bezeichnet  $H$  die Menge der *Hauptakteure* einer Szene, also all jene Agenten, die im Mittelpunkt der natürlichsprachlichen Beschreibung stehen sollen.  $E$  bezeichnet den Erzählstil ( $E \in \{Bericht, SBB, SBP\}$ ). Dabei ist zu beachten, daß bei einer szenischen Beschreibung aus Protagonistensicht (SBP, vgl. Abschnitt 2.3.3.3) nur ein Agens als Hauptakteur gewählt werden kann, nämlich der erzählende Protagonist.

### 3.4.2 Situationsbestimmung

Unabhängig vom ermittelten Erzählstil  $E$  wird für jeden Hauptakteur  $H \in H$  und für jeden Halbbildzeitpunkt, zu dem er in der betrachteten Szene aktiv ist, inkrementell die aktuelle Situation des Hauptakteurs bestimmt. Die Situationsbestimmung erfolgt anhand eines Situationsgraphenbaums, der üblicherweise zu beobachtende Abläufe an Straßenverkehrskreuzungen und Tankstellen repräsentiert. Durch Traversierung dieses Baumes wird für jeden Hauptakteur eine Folge besuchter Situationsknoten ermittelt. Die einzelnen Situationen werden durch ihren Situationsnamen bezeichnet, zu dem desweiteren die Bezeichnungen der einzelnen beteiligten Individuen angegeben sind. Diese Datenmenge bezeichnet ‘das Gleiche’ nach [Lindemann 87] (vgl. Abbildung 2.1). Bevor im nächsten Abschnitt darauf eingegangen wird, wie aus den berechneten Situationsfolgen perspektivische Beschreibungen abgeleitet werden, wird im folgenden zunächst auf den Situationsgraphenbaum, anhand dessen die Situationsbestimmung erfolgt, eingegangen. Dieser Situationsgraphenbaum ist in Anhang B abgedruckt (Situationsbestimmungsbaum). Er geht zurück auf [Haag 94], [Haag 95], [Theilmann 97] (vgl. auch [Kollnig et al. 95] und [Nagel et al. 95]), wurde aber im Rahmen der vorliegenden Arbeit stark erweitert. Sein Aufbau wird im folgenden beschrieben.

Die Auswertung beginnt zum initialen Zeitpunkt **1**. Ist das Agens zum darauffolgenden Zeitpunkt bereits aktiv, so wird in die Situation *sit\_aktiv* übergegangen. Andernfalls



des Tankplatzes (*sit\_leaving\_dispensing\_pump*) und schließlich das Verlassen des Tankstellengeländes (*sit\_leaving\_filling\_station*). Zum leichteren Verständnis der folgenden Ausführungen ist zur Erläuterung das verwendete Fahrbahnmodell angegeben (Abbildung 3.2). Beim Heranfahren an die Zapfsäule befindet sich das Agens in der Situation, daß es sich für eine bestimmte Zapfsäule entscheiden muß (*sit\_wishing\_to\_tank*). In der gewählten Modellierung wird davon ausgegangen, daß jedes Fahrzeug an jeder Zapfsäule den für sich relevanten Kraftstoff erhalten kann. Die Entscheidung, welche Zapfsäule das Agens letztendlich ansteuert, richtet sich demnach nach dem aktuellen Belegungszustand der Zapfsäulen. Folglich analysiert das Agens beim Heranfahren diesen aktuellen Belegungszustand für beide Tankspuren (*sit\_filling\_place\_occupation\_lower\_lane* bzw. *sit\_filling\_place\_occupation\_upper\_lane*). Hierbei besteht für beide Tankspuren die Möglichkeit, daß keine, ein, zwei oder alle drei Tankplätze belegt sind. Hat das Agens den aktuellen Belegungszustand analysiert, so wird es sich im folgenden für eine der beiden Tankspuren entscheiden (*sit\_choosing\_lower\_filling\_lane* bzw. *sit\_choosing\_upper\_filling\_lane*). Dann erfolgt die Wahl des Tankplatzes (*sit\_choosing\_dispensing\_pump*). Letztgenannte Situation wird wiederum in zahlreiche Untersituationen spezialisiert. Der weitere Ablauf hängt entschieden davon ab, ob der erste Tankplatz besetzt ist (*sit\_first\_filling\_place\_occupied*) oder nicht. Im negativen Fall ist zu untersuchen, ob der zweite Tankplatz besetzt ist (*sit\_second\_filling\_place\_occupied*). Sind beide Tankplätze frei, so wird das Agens sich zunächst einmal dem ersten Tankplatz nähern (*sit\_driving\_to\_first\_filling\_place*). Diesen wird es irgendwann einmal erreichen (*sit\_reaching\_first\_filling\_place*). Dann kann es zum zweiten Tankplatz weiterfahren und diesen erreichen, und dann gegebenenfalls auch zum letzten Tankplatz fahren und diesen erreichen (Situationsnamen analog). Ist hingegen der erste Tankplatz bereits besetzt (*sit\_first\_filling\_place\_occupied*), so hat das Agens zwei Möglichkeiten. Es kann erstens zum Überholen ansetzen, um den besetzten Tankplatz zu umfahren (*sit\_overtaking\_vehicle*), falls es sich auf einer Tankspur mit Ausweichspur befindet (obere Tankspur). Alternativ kann es sich dem besetzten Tankplatz nähern (*sit\_approaching\_first\_filling\_place*) und dann warten, bis der Tankplatz wieder frei ist (*sit\_waiting\_in\_front\_of\_first\_filling\_place*). Nach dem Überholen wird das Fahrzeug wieder von der Ausweichspur auf die Tankspur zurückkehren (*sit\_change\_back\_to\_filling\_lane*). Dann wird es zunächst den nächstfolgenden Tankplatz ansteuern (*sit\_driving\_to\_next\_dispensing\_pump*). Ein Sonderfall kann eintreten, wenn nur der zweite Tankplatz besetzt ist (*sit\_second\_filling\_place\_occupied*). Das Agens kann sich dazu entscheiden, nicht den bequem erreichbaren ersten Tankplatz zu besetzen, sondern statt dessen den zweiten Tankplatz zu umfahren und den letzten Tankplatz anzusteuern. Dann kann es passieren, daß das Agens beim Wiedereinschwenken von der Ausweichspur auf die Tankspur bereits über den letzten Tankplatz hinausfährt (*sit\_passing\_last\_filling\_place*). Es wird dann hinter dem letzten Tankplatz anhalten (*sit\_remain\_standing\_behind\_last\_filling\_place*) und dann zu diesem zurücksetzen (*sit\_backing\_up\_behind\_last\_filling\_place*). Hat das Fahrzeug endlich den anvisierten Tankplatz erreicht, so wird als nächstes das Fahrzeug betankt (*sit\_taking\_in\_petrol*). Diese Situation wird nicht weiter spezialisiert. Sie gilt als gegeben, wenn ein Agens auf einem Tankplatz stehen bleibt. Fährt das Agens schließlich

weiter, wird die Situation *sit\_leaving\_dispensing\_pump* ausgeprägt. Auch das Verlassen der Zapfsäule wird in Untersituationen spezialisiert. Es hängt vom Belegungszustand derjenigen Tankplätze ab, die sich zwischen dem eigenen Tankplatz und der Tankstellenausfahrt befinden. Befindet sich das Agens bereits auf dem letzten Tankplatz, so ist der Weg frei (*sit\_free\_exit*). Der Weg ist hingegen nicht frei (*sit\_no\_free\_exit*), wenn sich das Agens auf dem ersten Tankplatz befindet (*leaving\_first\_filling\_place*) und der zweite (*sit\_second\_filling\_place\_occupied2*) oder der dritte Tankplatz belegt ist (*sit\_last\_filling\_place\_occupied*). Ebenso ist der Weg nicht frei, wenn sich das Agens auf dem zweiten Tankplatz befindet (*sit\_leaving\_second\_filling\_place*) und der letzte Tankplatz belegt ist. Die Situation *sit\_no\_free\_exit* wird wie folgt spezialisiert. Befindet sich das Agens auf der unteren Tankspur, welche keine Ausweichspur besitzt, so kann es nur näher an den besetzten Tankplatz heranfahren (*sit\_approaching\_patiens\_when\_leaving*) und dann warten, bis dieser Tankplatz frei wird (*sit\_waiting\_behind\_patiens\_when\_leaving*). Andernfalls kann es den besetzten Tankplatz umfahren (*sit\_passing\_patiens\_when\_leaving*). Dazu muß es gegebenenfalls aus Platzmangel vorher ein Stück zurücksetzen (*sit\_backing\_up\_behind\_patiens\_when\_leaving*). Dann wird es einen Moment (zum Wechsel aus der Rückwärts- in die Vorwärtsbewegung) verharren (*sit\_waiting\_behind\_patiens\_when\_leaving*) und anschließend zum Überholen ansetzen (*sit\_driving\_off\_behind\_patiens\_when\_leaving*).

Bei einer Kreuzungsszene läßt sich das Überqueren einer Kreuzung (*sit\_crossing\_intersection1*) unterteilen in die Bestimmung der Herkunftsfahrbahn (*sit\_provenance*), das Heranfahren an die Kreuzung (*sit\_driving\_to\_intersection*), das eigentliche Überqueren (*sit\_crossing\_intersection*) und das Verlassen der Kreuzung (*sit\_leaving\_intersection*). Das Heranfahren an eine Kreuzung wird spezialisiert in eine Phase, in der man sich der Kreuzung nähert (*sit\_approach\_intersection*), und in eine Phase, in der man dort ggf. anhalten muß (*sit\_stopping\_at\_intersection*) und kurz wartet (*waiting\_at\_intersection*), bevor man weiterfahren kann (*starting\_at\_intersection*). Beim Heranfahren an die Kreuzung muß man sich, falls vorhanden, für eine Abbiegespur entscheiden (*sit\_taking\_left\_hand\_turn\_off\_lane*, *sit\_taking\_right\_hand\_turn\_off\_lane* bzw. *sit\_taking\_straight\_ahead\_lane*). Demzufolge wird auf der Kreuzung nach links, rechts oder gar nicht abgelenkt (*sit\_turning\_left*, *sit\_turning\_right* bzw. *sit\_driving\_straight\_ahead*). Sind die Namen der neu eingenommenen Straßen bekannt, so werden diese ebenfalls ausgeprägt (*sit\_turning\_left\_into* bzw. *sit\_turning\_right\_into*).

Wie bereits ausgeführt, werden im Situationsgraphenbaum *üblicherweise* zu beobachtende Abläufe modelliert. Es ist deshalb interessant zu fragen, was passiert, wenn sich ein Agens nicht so verhält, wie es üblich ist. Solange sein Verhalten der Modellierung entspricht, ist zu erwarten, daß zu jedem Zeitpunkt ein speziellster Situationsknoten des Baumes ausgeprägt werden kann. Verhält sich dann das Agens plötzlich anders als vorgesehen, so bricht der Auswertungsprozeß deswegen nicht zusammen. Statt dessen wird das inkonsistente Verhalten des Agens dadurch deutlich, daß nur noch allgemeinere Situationsknoten ausgeprägt werden können. Die Reaktion des Systems darauf, beziehungsweise die Behandlung dieser Verhaltensinkonsistenz bei der Generierung des



natürlichsprachlichen Textes unterliegt dann dem nächstfolgenden Auswertungsschritt, in dem aus der ermittelten Situationsfolge eine perspektivische begriffliche Beschreibung abgeleitet wird. Somit wird das Problem ‘inkonsistenten Agensverhaltens’ erst im nächsten Abschnitt interessant.

Für die oben angegebenen Situationsnamen wurden Bezeichnungen gewählt, die auf den Leser intuitiv wirken. Die Semantik dieser Bezeichnungen spielt rechnerintern bei der weiteren Auswertung keine Rolle. So ist es für den weiteren Auswertungsprozeß ohne Belang, ob eine Situation intuitiv als *sit\_wishing\_to\_tank* oder weniger intuitiv als *sit\_asdf1234* bezeichnet wird, solange diese Bezeichnungen fest und eindeutig sind. Damit enthalten die Situationsnamen zu diesem Zeitpunkt der Auswertung keinerlei perspektivische Interpretation. Diese wird erst im nächsten Auswertungsschritt, beschrieben im nächsten Abschnitt, geleistet. Dort wird anhand des ermittelten Erzählerstandpunktes entschieden, ob beispielsweise die Situation *sit\_wishing\_to\_tank* vorsichtig als möglicher Beginn eines Tankvorgangs gedeutet wird, oder optimistisch-prognostisch als sicherer Beginn eines Tankvorgangs, oder gar lediglich die Tatsache formuliert wird, daß ein Fahrzeug auf ein Tankstellengelände gefahren ist, ohne über eine mögliche Absicht des Fahrers zu spekulieren.

### 3.4.3 Perspektivische Beschreibungen von Situationsfolgen

Im vorangegangenen Abschnitt wurde beschrieben, wie für jeden Hauptakteur die Folge von Situationen bestimmt wird, in denen er sich nach und nach befindet. Jetzt erfolgt die Ableitung einer perspektivischen, begrifflichen Beschreibung aus dieser Situationsfolge. Die Perspektive, mit der die gegebene Situationsfolge betrachtet wird, ergibt sich unmittelbar aus dem zuvor bestimmten Erzählerstandpunkt. Dabei ist es Teil des Weltwissens des Systems, wie es bestimmte Situationsfolgen in einer bestimmten Perspektive zu beschreiben hat. Dieses Wissen wird in Form von Situationsgraphen repräsentiert. Es sind grundsätzlich drei Varianten zu berücksichtigen, wie Situationen perspektivisch beschrieben werden (vgl. Abbildung 3.3).

- Der Regelfall besteht darin, daß eine Situation mit Prädikaten beschrieben wird, die zeitlich auf derselben Ebene liegen, d.h. demselben Gültigkeitsintervall zugeordnet werden können. Will man z.B. die Situation begrifflich beschreiben, daß ein Fahrzeug auf einem Tankplatz hält, um zu tanken, so kann man dies mit Hilfe von drei Prädikaten tun. Das erste repräsentiert das beobachtete Geschehen (anhaltend), das zweite die verfolgte Absicht (tanken), und das letzte stellt den kausalen Zusammenhang zwischen beiden Beobachtungen her. Alle drei Prädikate können ohne Verletzung kontextueller, physikalischer oder sonstiger Gesetze demselben Gültigkeitszeitraum zugeordnet werden. Damit liegen sie auf einer zeitlichen Ebene. Die perspektivische Beschreibung von Situationen ausschließlich auf einer zeitlichen Ebene wird im folgenden als Regelfall betrachtet. Er korrespondiert zu den Fällen  $k_1 = 1, k_2 = 1, \dots, k_n = 1$  in Abbildung 3.3.

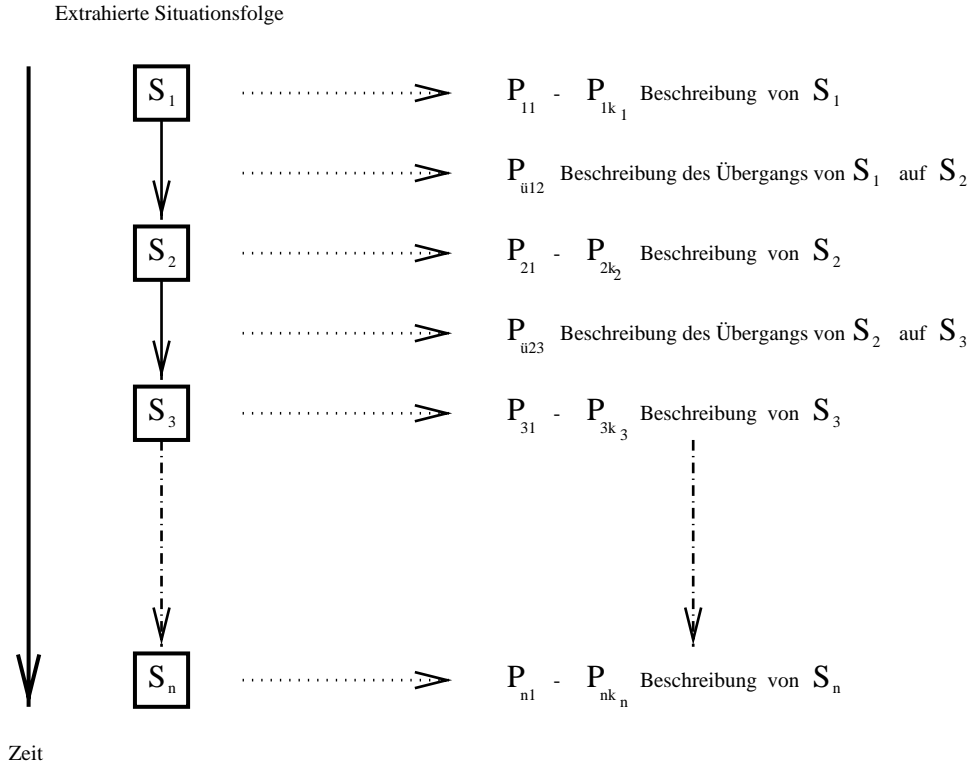


Abbildung 3.3: Varianten zur Ableitung perspektivischer Beschreibungen einer extrahierten Situationsfolge. Links in der Abbildung ist die extrahierte Situationsfolge in ihrem zeitlichen Verlauf angegeben, rechts die ableitbaren perspektivischen Beschreibungen.

- Die Fälle  $k_i > 1$  stehen für perspektivische Beschreibungen auf unterschiedlichen zeitlichen Ebenen. So kann beispielsweise das Anhalten, um zu tanken, auch als zwei Geschehen gewertet werden, die zwar zeitlich aufeinanderfolgen, aber derselben Situation zugeordnet werden (erst anhalten, dann tanken). Dies ist im Beispielfall deshalb gegeben, weil das Stehen auf dem Tankplatz systemintern bisher als sicheres Indiz für den Beginn eines Betankungsvorgangs gewertet wird. Insbesondere findet bisher keine Personenverfolgung statt, die ermittelt, ob der Fahrer tatsächlich aussteigt und den Wagen betankt. Der Betankungsvorgang kann folglich durch zwei Geschehen beschrieben werden, die zwar derselben Situation zugeordnet werden, die aber untereinander nicht dasselbe Gültigkeitsintervall besitzen.
- Die dritte zu berücksichtigende Beschreibungsvariante besteht darin, daß man nicht die Situation selbst, sondern den Übergang von einer Situation auf eine andere beschreibt (in Abbildung 3.3 gekennzeichnet mit  $P_{iab}$ ).

Die Ableitung der perspektivischen Beschreibung  $\mathbf{P}$  aus der Situationsfolge  $\mathbf{S}$  erfolgt durch Situationsgraphentraversierung. Das bedeutet, daß ein Situationsgraphenbaum

benötigt wird, der aus einer zuvor extrahierten Situationsfolge  $\mathbf{S}$  die perspektivische Beschreibung  $\mathbf{P} = \{P_{ij}, P_{üab} : 1 \leq i \leq n, 1 \leq j, b = a + 1, 1 \leq a \leq n - 1\}$  in Abhängigkeit von der extrahierten Situationsfolge berechnet. Situationsgraphen, die entsprechendes leisten, werden im folgenden als *Perspektivierungsgraphen* bezeichnet. Man kann es grundsätzlich dem Systementwerfer überlassen, Perspektivierungsgraphen zu entwerfen. Dabei ist jedoch zu bedenken, daß für jede mögliche Perspektive  $\mathbf{P}$  aus der Menge aller denkbaren Perspektiven  $\mathbf{P}$  ein spezieller Perspektivierungsgraph vonnöten ist. Die Modellierung dieser Graphen ist zudem aufwendig und schwierig. Deshalb wurde im Rahmen dieser Arbeit das Programm MALCOLM entwickelt, das automatisch aus einem gegebenen Situationsbestimmungsbaum und einer gegebenen Perspektive  $\mathbf{P}$  den zugehörigen Perspektivierungsbaum berechnet. Das Wirkungsszenario von MALCOLM ist in Abbildung 3.4 grafisch veranschaulicht. Gegeben sind ein durch Benutzer- oder Spezifikationsvorgaben eindeutig festgelegter Erzählerstandpunkt und ein Situationsbestimmungsbaum, mit dessen Hilfe aus den extrahierten Fakten der Faktendatenbank die Folge besuchter Situationsknoten berechnet wird. Aus dem Erzählerstandpunkt läßt sich die Perspektive  $\mathbf{P}$  ableiten, mit Hilfe derer die Situationsfolge beschrieben werden soll. Für  $\mathbf{P}$  ist als Teil des systeminternen Wissens in tabellarischer Form angegeben, wie Situationen oder Situationsübergänge mit dieser Perspektive beschrieben werden sollen. Das Programm MALCOLM erzeugt aus diesen Tabellen und aus dem Situationsbestimmungsbaum den die ermittelte Perspektive  $\mathbf{P}$  repräsentierenden Perspektivierungsbaum, mit Hilfe dessen aus der extrahierten Situationsfolge die entsprechende perspektivische Beschreibung abgeleitet werden kann.

### 3.4.4 Automatisierung

In Anhang B ist der in dieser Arbeit verwendete Situationsbestimmungsbaum angegeben. Das Zustandsschema jeder Situation enthält diejenigen Prädikate, die ausgeprägt werden müssen, um ein Agens mit dieser Situation zu assoziieren. Im Falle der erfolgreichen Ausprägung wird der Situationsname und der Name der beteiligten Agenten und Bezugsobjekte ausgegeben. Diese Ausgabe erfolgt in der bereits bekannten Notation unscharf metrisch temporaler Logikfakten und wird deshalb als neues Faktum in die Faktendatenbank eingetragen. Die neuen Fakten bilden die Datenmenge, anhand derer der von MALCOLM aufgebaute Perspektivierungsbaum ausgeprägt wird. Dabei sind, wie bereits im vorigen Abschnitt angedeutet, drei Varianten denkbar (siehe Abbildung 3.3 und Abbildung 3.5). In letztgenannter Abbildung ist oben die Umsetzung einer perspektivischen Beschreibung der Variante **1** gezeigt. Wird bei der Situationsbestimmung die Situation ‘*ErsteSituation*’ ausgeprägt, so wird mit Hilfe des **note**-Prädikates ein neues Faktum in der Faktendatenbank eingetragen, dessen Prädikatsname dem Situationsnamen entspricht (links oben in der Abbildung). Anhand dieses Faktums wird bei der Traversierung des Perspektivierungsgraphen erkannt, ob die entsprechende Situation bei der bereits abgeschlossenen Situationsbestimmung auftrat oder nicht. Dazu muß das Zustandsschema der entsprechenden Situation im Perspektivierungs-

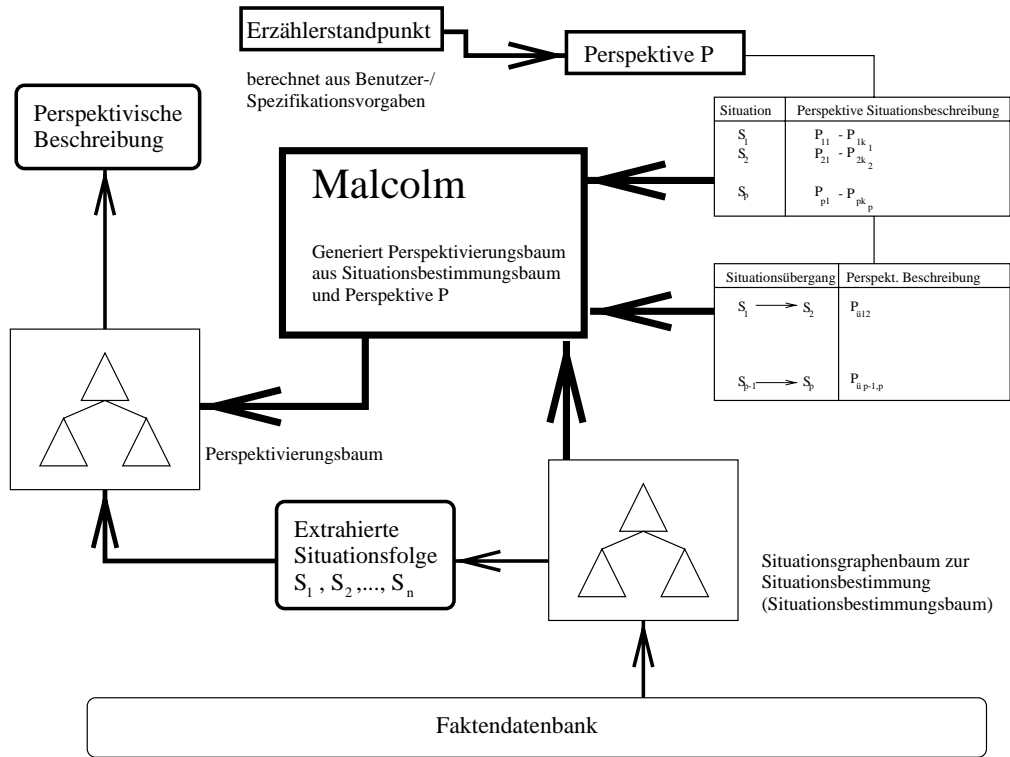


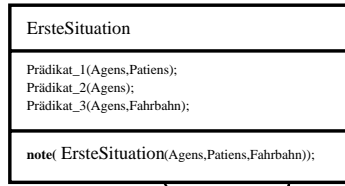
Abbildung 3.4: Einbettung und Aufgabe des Programms MALCOLM bei der Berechnung perspektivischer Beschreibungen. Nähere Erläuterungen im Text.

baum genau das Prädikat enthalten, das bei der vormaligen Situationsbestimmung in die Faktendatenbank eingetragen wurde (rechts oben in der Abbildung). Mit Hilfe des **note**-Prädikats wird dann diejenige perspektivische Beschreibung ausgegeben, die in der Perspektivierungstabelle vorgegeben ist.

Die mittlere Zeile der Abbildung 3.5 zeigt die Realisierung der Beschreibung des Übergangs von der Situation 'Erste Situation' auf die Situation 'Zweite Situation' des Situationsbestimmungsbaums. Dazu wird im Perspektivierungsbaum eine eigene Situation vorgesehen, deren Namen sich aus besagten beiden Situationsnamen verbunden mit der Zeichenfolge '\_to\_' ergibt. Das Zustandsschema dieser Situation entspricht demjenigen Prädikat, das bei der Assoziation von 'Zweite Situation' ausgegeben wurde. So wird erreicht, daß zum gleichen Auswertungszeitpunkt, an dem der Übergang von 'Erste Situation' auf 'Zweite Situation' bei der Situationsbestimmung erfolgte, auch der Übergang von 'Erste Situation' auf 'Erste Situation\_to\_Zweite Situation' erfolgt.

Die untere Zeile von Abbildung 3.5 zeigt die Realisierung der perspektivischen Beschreibung der Situation 'Zweite Situation' (links), die, laut Eintrag in der Perspektivierungstabelle (Mitte), in zwei Schritten erfolgen soll. Dadurch werden im Perspektivierungsgraphen (rechts) zwei Situationen vorgesehen. Erstere realisiert den ersten Teil der Beschreibung, letztere den zweiten. Beide Situationen besitzen dasselbe Zustands-

Situationsbestimmungsbaum



Perspektivierungstabelle

ErsteSituation
1   persp_praedikat_1(Agens); persp_praedikat_2(Agens,Fahrbahn); persp_praedikat_3(Agens,Patiens);

ErsteSituation_to_ZweiteSituation
1   persp_praedikat_4(Agens); persp_praedikat_5(Agens,Fahrbahn); persp_praedikat_6(Agens,Patiens);

ZweiteSituation
1   persp_praedikat_7(Agens); persp_praedikat_8(Agens,Fahrbahn);
2   persp_praedikat_9(Agens,Patiens);

Perspektivierungsbaum

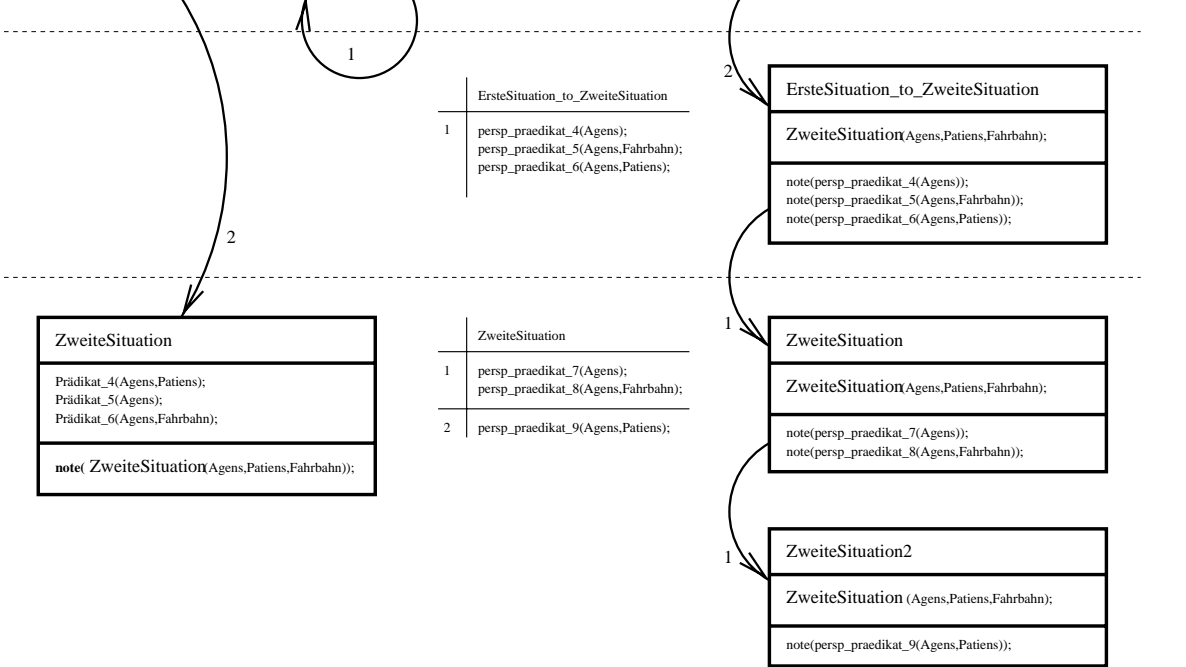
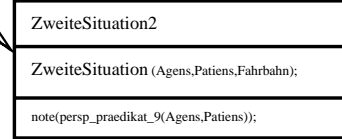
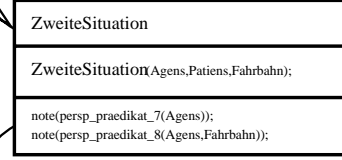
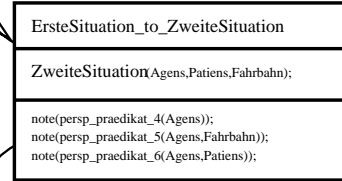
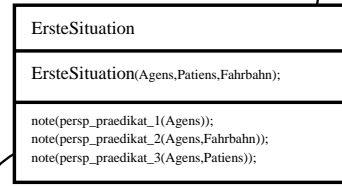


Abbildung 3.5: **Links:** Gegeben: Zwei Situationen des Situationsbestimmungsbaumes. **Mitte:** Einträge in die Perspektivierungstabelle bezüglich der perspektivischen Beschreibung dieser beiden Situationen und deren Übergang. **Rechts:** Resultierender Situationsgraph des Perspektivierungsbaumes. Nähere Erläuterung im Text.

schema, welches der Ausgabe bei der Assoziation von ‘*ZweiteSituation*’ im Perspektivierungsbaum entspricht. Ihr Name unterscheidet sich durch das Anhängen der Ziffer ‘2’ an den Namen der zweiten Situation.

Die in Abbildung 3.5 gezeigten beiden Situationen des Situationsbestimmungsbaums (links) besitzen folgende, übliche Nachfolgeregelung: Erster potentieller Nachfolger der Situation ‘*ErsteSituation*’ ist die Situation selbst, zweiter potentieller Nachfolger ist ‘*ZweiteSituation*’. Diese beiden Situationen und ihr Übergang werden wie oben ausgeführt durch den rechts in der Abbildung angegebenen Ausschnitt aus einem Perspektivierungsbaum perspektivisch beschrieben. Dabei wurde bisher noch nicht auf die Nachfolgeregelungen der vier Situationen des Perspektivierungsbaums eingegangen. Die erste angegebene Situation besitzt den gleichen Namen und die gleiche Nachfolgeregelung wie die entsprechende erste Situation des Situationsbestimmungsbaums. Allerdings ist ihre zweite potentielle Nachfolgesituation die Situation

‘*ErsteSituation\_to\_ZweiteSituation*’, welche den Übergang im Situationsbestimmungsbaum beschreibt. Diese Situation besitzt dasselbe Zustandsschema wie die beiden Situationen ‘*ZweiteSituation*’ und ‘*ZweiteSituation2*’. Um zu erreichen, daß alle diese Situationen ausgeprägt werden können (und dies in der logisch richtigen Reihenfolge), werden die Situationen ‘*ErsteSituation\_to\_ZweiteSituation*’ und ‘*ZweiteSituation*’ nicht als ihre eigenen Nachfolger modelliert (siehe Abbildung 3.5). So wird erreicht, daß alle vier Situationen in der intendierten Weise traversiert werden. Die im Rahmen dieser Arbeit entwickelte Systemkomponente MALCOLM berechnet automatisch aus einem gegebenen Situationsbestimmungsbaum und einer Perspektivierungstabelle den entsprechenden Perspektivierungsbaum. Die Funktionsweise von MALCOLM ist im Anhang detailliert dargestellt.

## 3.5 Erweiterte Schlußfolgerungen

In den vorangegangenen Abschnitten wurde gezeigt, wie perspektivische Beschreibungen von Fahrzeugverhalten algorithmisch erstellt werden. Im Mittelpunkt einer solchen Beschreibung steht immer genau ein Hauptakteur. Soll jedoch eine Beschreibung erstellt werden, die mehr als einen ( $n > 1$ ) Hauptakteur kennt, so wird dies bisher dadurch realisiert, daß für jeden der  $n$  Hauptakteure eine Einzelbeschreibung berechnet wird. Die  $n$  Einzelbeschreibungen werden dann zu einer gruppierenden Beschreibung zusammengefügt, die alle  $n$  Hauptakteure gleichermaßen berücksichtigt. Bei der Zusammenfügung von mehreren Einzelbeschreibungen zu einer gruppierenden Beschreibung ergeben sich oft weitere Schlußfolgerungs- und Abstraktionsmöglichkeiten, die die Aussagekraft erhöhen. Zwei Arten solcher Schlußfolgerungs- und Abstraktionsmöglichkeiten wurden in dieser Arbeit berücksichtigt:

1. Ausnutzung von zusätzlichem Kontextwissen zur Ableitung erweiterter Schlußfolgerungen.
2. Ableitung quantorsierter Aussagen.

### 3.5.1 Ausnutzung von zusätzlichem Kontextwissen

Bei der Zusammenführung mehrerer Einzelbeschreibungen zu einer gruppierenden Beschreibung ergeben sich oft weitere Schlußfolgerungsmöglichkeiten. Dies sei anhand eines Beispiels demonstriert.

Abbildung 1.1 zeigt die automatisch berechneten Trajektorien der beiden Fahrzeuge ‘*obj\_9*’ und ‘*obj\_4*’. Abbildung 3.6 zeigt automatisch berechnete begriffliche, perspektivische Beschreibungen des Verhaltens der beiden Fahrzeuge. Diese beiden Beschreibungen wurden unabhängig voneinander erstellt. Das Fahrzeug ‘*obj\_9*’ muß beim Verlassen der Zapfsäule ein anderes Fahrzeug umfahren, das sich in der Zwischenzeit vor es gestellt hat. Dieses Fahrzeug wird in der Beschreibung schlicht als ‘*other\_car*’ bezeichnet,

(9.1)	300	:	300	!	enter(obj_9, the_filling_station, 1).
(9.2)	305	:	305	!	drive_on(obj_9, the_upper_filling_lane, 3).
(9.3)	409	:	409	!	stop_on(obj_9, the_second_filling_place, 17).
(9.4)	410	:	1838	!	take_in(obj_9, petrol, 19).
(9.5)	1840	:	1840	!	stand_on(other_car, the_last_filling_place, 21).
(9.6)	1841	:	1958	!	reason(24, 22, 23).
(9.7)	1841	:	1958	!	certainty(25, 23, presumably).
(9.8)	1841	:	1958	!	back_up(obj_9, 22).
(9.9)	1841	:	1958	!	want_to_overtake(obj_9, other_car, 23).
(9.10)	2155	:	2155	!	change_to(obj_9, the_passing_lane, 26).
(9.11)	2156	:	2318	!	pass(obj_9, other_car, 27).
(9.12)	2319	:	2320	!	leave(obj_9, the_filling_station, 20).
(4.1)	614	:	614	!	enter(obj_4, the_filling_station, 1).
(4.2)	619	:	619	!	drive_on(obj_4, the_upper_filling_lane, 3).
(4.3)	620	:	620	!	stand_on(other_car, the_second_filling_place, 4).
(4.4)	627	:	627	!	change_to(obj_4, the_passing_lane, 11).
(4.5)	628	:	881	!	overtake(obj_4, other_car, 12).
(4.6)	882	:	882	!	change_back_to(obj_4, the_filling_lane, 13).
(4.7)	984	:	1071	!	stop_behind(obj_4, the_last_filling_place, 14).
(4.8)	1072	:	1165	!	back_up_to(obj_4, the_last_filling_place, 15).
(4.9)	1227	:	1227	!	stop_on(obj_4, the_last_filling_place, 18).
(4.10)	1228	:	4707	!	take_in(obj_4, petrol, 19).
(4.11)	4708	:	4898	!	leave(obj_4, the_filling_station, 20).

Abbildung 3.6: Automatisch berechnete perspektivische Beschreibung für das Verhalten der Fahrzeuge ‘obj\_9’ und ‘obj\_4’ der Tankstellenszene (vgl. Abbildung 1.1).

da keine Veranlassung bestand, ihm einen eigenen Namen zu geben. Analog umfährt das Fahrzeug ‘obj\_4’ auf seinem Weg zum letzten Tankplatz ebenfalls ein ‘anderes Auto’. Ein Blick auf Abbildung 1.1 zeigt, daß es sich bei dem ‘anderen Auto’ in der Beschreibung zu ‘obj\_4’ um das Fahrzeug ‘obj\_9’ handelt, während in der Beschreibung zu ‘obj\_9’ das Fahrzeug ‘obj\_4’ das ‘andere’ ist. Es stellt sich die Frage, wie bei der Zusammenfügung der beiden Beschreibungen aus Abbildung 3.6 eben jene Schlußfolgerungen mit Hilfe der metrisch temporalen Logik erschlossen werden können. Das folgende Programm ist eine mögliche Lösung:

```
always (overtake2(A,B,X) :- overtake(A,other_car,X) ,
take_in(B,petrol,X2) , standing_on(B,FB,TP) ,
```

```
formerly_driving_on_FB_close_to_TP(A,FB,TP) ,  
note(overtake(A,B,X))).
```

```
always (standing_on(B,FB,TP) :- event (drive_on(B,FB,X3) ,  
a : e ? stop_on(B,TP,X4))).
```

```
always (formerly_driving_on_FB_close_to_TP(A,FB,TP) :-  
event (drive_on(A,FB,X3) ,  
next stand_on(other_car,TP,X5))).
```

Dieses Programm enthält die temporallogische Entsprechung der folgenden allgemeingültigen natürlichsprachlichen Feststellung: *Wenn ein Fahrzeug A ein anderes Fahrzeug überholt, während ein Fahrzeug B auf einem bestimmten Tankplatz TP einer bestimmten Tankspur FB steht und tankt, und wenn Fahrzeug A zudem sich unmittelbar vor Beginn des Überholvorgangs hinter besagtem Fahrzeug B befand, dann handelt es sich bei dem überholten Fahrzeug um Fahrzeug B.* Einzelheiten zur Syntax und Semantik des Programms entnehme man [Schäfer 96].

Metrisch temporale Logikprogramme dieser Art bieten bequem die Möglichkeit, kontextuelles Weltwissen in die Auswertung mit einzubringen. Das oben angegebene Programm erlaubt, Faktum 4.5 aus Abbildung 3.6 zu ersetzen durch das aussagekräftigere Faktum

(4.5') 628 : 881 ! overtake(obj\_4, obj\_9, 12).

### 3.5.2 Quantorisierte Geschehen

Im vorigen Abschnitt wurde gezeigt, wie man bei der Zusammenführung mehrerer Einzelbeschreibungen zu einer gruppierenden Beschreibung erweiterte Schlußfolgerungen treffen kann. Wenn die Zahl der Einzelbeschreibungen sehr hoch ist, kann es jedoch auch zweckmäßig sein, nicht noch weitere Informationen aus der Datenmenge abzuleiten, sondern die vorhandenen Daten zum Zwecke einer kompakteren Beschreibung weiter zu abstrahieren. Gruppiert man beispielsweise einhundert Einzelbeschreibungen zusammen, in denen jeweils ein Fahrzeug beim Linksabbiegen auf einer Kreuzung beobachtet wird, dann ist eine folgende Beschreibung eher unangebracht:

*Objekt\_1 bog links ab. Dann bog Objekt\_2 links ab. Dann bog auch noch Objekt\_3 links ab. Dann bog Objekt\_4 ... Als letztes bog Objekt\_100 links ab.*

Ohne wesentlichen Verlust an Information läßt sich ein solcher Ablauf mit den Worten 'Einhundert Fahrzeuge bogen links ab' oder 'Viele Fahrzeuge bogen links ab' prägnanter beschreiben. Die Zusammenfassung von Einzelgeschehen zu solchen quantorisierten Geschehen ist ein wesentlicher Ansatzpunkt dieser Arbeit und wird im folgenden behandelt. Dabei wird unterschieden zwischen *Zahlwörtern* und *unscharfen Quantoren* wie z.B. *viele*, *einige*, *wenige*.



### 3.5.2.1 Unscharfe Quantoren

Unter quantorisierten Beschreibungen werden solche Beschreibungen verstanden, welche vergleichbares Verhalten verschiedener Agenten in einer Aussage unter Verwendung eines natürlichsprachlichen Quantors versprachlichen. Ein typisches quantorisiertes Geschehen beinhaltet beispielsweise die Aussage: *Einige Fahrzeuge biegen links ab*. Eine solche Aussage ist im Prinzip mehrdeutig in ihrer *Lesart* und in der *Definition des zugrundeliegenden Quantors*. Hinsichtlich des erstgenannten Kriteriums lassen sich die *universelle* und die *existentielle* Lesart unterscheiden.

- **Universelle Lesart:** Die Aussage wird so gedeutet, daß zu *jedem* Zeitpunkt des Gültigkeitszeitraums einige Fahrzeuge links abbiegen, wobei dies immer wieder andere Fahrzeuge sein können. Daraus kann in der Summe resultieren, daß insgesamt alle Fahrzeuge links abbiegen, aber immer nur einige zur gleichen Zeit.
- **Existenzielle Lesart:** Die Aussage wird so gedeutet, daß *insgesamt* einige Fahrzeuge links abbiegen, wobei es durchaus Teilintervalle des Gültigkeitszeitraums geben kann, in denen beispielsweise kein einziges Fahrzeug links abbiegt. Dennoch bleibt die Aussage wahr.

Es ist festzustellen, daß die existentielle Lesart von quantorisierten Aussagen diejenige ist, die dem gängigen Sprachgebrauch entspricht.

Zweites Kriterium zur unterschiedlichen Interpretation einer quantorisierten Geschehensbeschreibung ist die Definition des zugrundeliegenden Quantors. Hierbei lassen sich die grundmengenabhängige und grundmengenunabhängige Definition von Quantoren unterscheiden.

- **Grundmengenunabhängige Definition:** Die Zuordnung der Quantoren erfolgt allein abhängig von der absoluten Größe der betrachteten Menge: *200 Fahrzeuge sind einige, zwei Fahrzeuge sind nicht einige*. (Nebenbei sei erwähnt, daß selbst der grundmengenunabhängige Gebrauch natürlichsprachlicher unscharfer Quantoren uneinheitlich ist. So kann er beispielsweise von besonderen Eigenschaften der Individuen, über die quantorisiert wird, abhängen. Zum Beispiel sind 200 Menschen in einem Raum viele, 200 Bakterien im selben Raum aber eher wenige. Dies liegt letztendlich daran, daß auf dieser Welt im allgemeinen und in Räumen im besonderen üblicherweise viel mehr Bakterien vorhanden sind als Menschen. Solche kontextuellen Abhängigkeiten bei der Quantordefinition werden im folgenden jedoch nicht betrachtet, da grundsätzlich nur Fahrzeugverhalten beschrieben und folglich nur über Fahrzeugmengen quantorisiert wird).
- Bei der **grundmengenabhängigen Definition** liegt der Quantorzuordnung eine bestimmte Grundmenge zugrunde: *zwei von drei Fahrzeugen sind einige, 200 von 10.000 Fahrzeugen sind nicht einige*.

```

always card([], 0.0).

always (card([H/T], N1) :- card(T, N), N1 is N + 1.0).

always (portion(Vars,Domain,Body,Portion) :- call( findall List1 Vars Domain ),
call( findall List2 Vars (Domain , Body) ), card(List1, Card1), card(List2, Card2),
Portion is Card2 / Card1).

always (all(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion ≥ 1.0).

always (nearly_all(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion ≥ 0.7 , Portion < 1.0).

always (most(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion > 0.5 , Portion < 1.0).

always (many(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion ≥ 0.4 , Portion < 0.8).

always (several(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion > 0.2 , Portion < 0.6).

always (some(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion > 0.1 , Portion < 0.5).

always (a_few(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion > 0.0 , Portion < 0.3).

always (no(Vars,Domain,Body)
        :- portion(Vars,Domain,Body,Portion), Portion = 0.0).

```

Abbildung 3.7: Grundmengenabhängige Modellierung natürlichsprachlicher Quantoren als unscharf metrisch temporale Logikprädikate. Entscheidend ist die Berechnung des Prädikats  $portion(Vars, Domain, Body, Portion)$ . Im ersten Schritt wird die Liste  $List1$  derjenigen Individuen aus  $Vars$  berechnet, die einer Grundmenge  $Domain$  angehören. Dann wird die Liste  $List2$  derjenigen Individuen aus  $Vars$  zusammengestellt, die sowohl der Grundmenge  $Domain$  angehören als auch dasjenige Prädikat erfüllen, über das quantorisiert werden soll ( $Body$ ). Der freie Bezeichner  $Portion$  enthält dann das Verhältnis der Anzahl der Individuen aus  $List2$  (berechnet in  $Card2$ ) zu der Anzahl der Individuen aus  $List1$  (berechnet in  $Card1$ ). Diese Verhältniszahl wird den einzelnen Prädikaten zugeordnet, die natürlichsprachliche Quantoren repräsentieren.

Die meisten Quantoren lassen sich sowohl grundmengenabhängig als auch -unabhängig verwenden. Welche Art der Verwendung im Einzelfall gerade gewählt wurde, wird meist

aus dem Kontext deutlich. Im täglichen Sprachgebrauch ist die Unterscheidung der beiden Quantordefinitionsarten meist unproblematisch, im Rahmen der begrifflichen Auswertung von Bildfolgen jedoch nicht, da beide Definitionsarten eine vollkommen unterschiedliche Berechnung der jeweiligen Quantoren erfordern. Abschließend sei noch darauf hingewiesen, daß bestimmte Quantoren generell nur grundmengenabhängig definiert werden können. Dies sind beispielsweise die Quantoren *alle*, *die meisten* und *die wenigsten*.

[Früauf 94] entwickelte eine ausschließlich grundmengenabhängige Repräsentation von natürlichsprachlichen Quantoren, basierend auf binären Wahrheitstabellen. Darauf aufbauend berechnen [Gerber&Nagel 96b] quantorisierte Geschehen universeller Lesart, aus denen jedoch durch ein stark heuristisches Verfahren quantorisierte Geschehen existentieller Lesart abgeleitet werden konnten. [Theilmann 97] führte erstmals ein Experiment zur Berechnung quantorisierter Beschreibungen mit Hilfe von F-LIMETTE durch, basierend auf einer ausschließlich grundmengenabhängigen Repräsentation natürlichsprachlicher Quantoren nach [Schäfer 96]. [Theilmann 97] berücksichtigte in seinen Experimenten lediglich die Berechnung quantorisierter Beschreibungen nach universeller Lesart, nicht aber nach existentieller Lesart, wie sie der gängige Sprachgebrauch verlangt.

Der Unterschied zwischen universeller und existentieller Lesart von quantorisierten Geschehen soll anhand des folgenden Gedankenexperiments verdeutlicht werden. Gesetzt den Fall, in einer Szene befinden sich zehn Fahrzeuge, die nacheinander links abbiegen, wobei zu jedem Zeitpunkt genau ein Fahrzeug links abbiegt. Dies sei durch folgende begriffliche Beschreibung dokumentiert:

1 : 10 ! linksabbiegen(obj\_1).  
 11 : 20 ! linksabbiegen(obj\_2).  
 21 : 30 ! linksabbiegen(obj\_3).  
 31 : 40 ! linksabbiegen(obj\_4).  
 41 : 50 ! linksabbiegen(obj\_5).  
 51 : 60 ! linksabbiegen(obj\_6).  
 61 : 70 ! linksabbiegen(obj\_7).  
 71 : 80 ! linksabbiegen(obj\_8).  
 81 : 90 ! linksabbiegen(obj\_9).  
 91 : 100 ! linksabbiegen(obj\_10).

Ein zeitpunktweise fortschreitender Algorithmus zur Berechnung quantorisierter Beschreibungen nach universeller Lesart erkennt, daß zu jedem Zeitpunkt genau ein Fahrzeug links abbiegt, und assoziiert zum Mengenverhältnis eins zu zehn den Quantor *wenige*. Daraus resultiert folgende quantorisierte Beschreibung:

```

1 : 10 ! linksabbiegen(wenige_Fahrzeuge).
11 : 20 ! linksabbiegen(wenige_Fahrzeuge).
21 : 30 ! linksabbiegen(wenige_Fahrzeuge).
31 : 40 ! linksabbiegen(wenige_Fahrzeuge).
41 : 50 ! linksabbiegen(wenige_Fahrzeuge).
51 : 60 ! linksabbiegen(wenige_Fahrzeuge).
61 : 70 ! linksabbiegen(wenige_Fahrzeuge).
71 : 80 ! linksabbiegen(wenige_Fahrzeuge).
81 : 90 ! linksabbiegen(wenige_Fahrzeuge).
91 : 100 ! linksabbiegen(wenige_Fahrzeuge).

```

Die Zusammenfassung der Gültigkeitsintervalle ergibt dann folgende Gesamtaussage:

```

1 : 100 ! linksabbiegen(wenige_Fahrzeuge).

```

Nach universeller Lesart ist dies eine richtige Aussage, denn zu jedem Zeitpunkt aus dem Intervall [1:100] gilt: Wenige Fahrzeuge sind zu diesem Zeitpunkt links abgebogen. Tatsächlich sind jedoch im Zeitintervall [1:100] nicht *wenige*, sondern **alle** Fahrzeuge links abgebogen. Diese Aussage entspricht dem existentiellen Gebrauch natürlichsprachlicher Quantoren und damit auch dem gängigen Sprachgebrauch. Entscheidend bei der Berechnung quantorisierte Beschreibungen ist folglich nicht, was genau an jedem einzelnen Zeitpunkt des betrachteten Zeitintervalls passiert, sondern vielmehr, was sich im Betrachtungszeitintervall *in der Summe* ereignet. Weder der Ansatz von [Früauf 94] noch derjenige von [Theilmann 97] leisten diese Art der Berechnung.

Die Herausforderung besteht folglich darin, die Möglichkeiten, die die unscharf metrisch temporale Logik bietet, zu nutzen, um quantorisierte Beschreibungen existentieller Art berechnen zu können. Hierzu wurde die in Abbildung 3.7 gezeigte Modellierung natürlichsprachlicher Quantoren gewählt. Die Quantoren sind unscharf definiert, so daß sich die Anwendbarkeitsbereiche zwischen ‘benachbarten’ Quantoren überlappen, wodurch geringfügige Schwankungen im prozentualen Anteil nicht zwangsläufig zu einem Alternieren zwischen verschiedenen Quantoren führen müssen. Das folgende metrisch temporale Logikprogramm realisiert die Berechnung quantorisierte Beschreibungen von Abbiegevorgängen in existentieller Lesart unter Verwendung der Quantorendefinition aus Abbildung 3.7.

```

always (quantorbestimmung(Vars, Domain, Body, Body2, Body3, Body4) :-
    quantorisierung(Vars, Domain, Body) ,
    quantorisierung(Vars, Body, Body2) ,
    quantorisierung(Vars, Body, Body3) ,
    quantorisierung(Vars, Body, Body4)).

```

```

always (quantorisierung(Vars, Domain, Body) :-
    call({ findall List1 Vars Domain }),
    call({ findall List2 Vars (Domain , Body) }),
    card(List1, Card1), card(List2, Card2),
    ((Card2 < 11 , zahlenname(Card2,V) ,
    note(quantor(V,Domain,Body))) ;
    (Card2 > 10 , Portion is Card2 / Card1 ,
    quantorname(Portion,W), note(quantor(W,Domain,Body))))).

always (on_intersection(X) :- event come_from(X,Y,Z)).

always (come_from(X,Y) :- event come_from(X,Y,Z)).

always (turn_right_into(X,Y) :- event turn_right_into(X,Y,Z)).

always (turn_left_into(X,Y) :- event turn_left_into(X,Y,Z)).

always (drive_straight_ahead(X) :- event drive_straight_ahead(X,Z)).

?- quantorbestimmung(X, on_intersection(X), come_from(X,Y1),
    turn_right_into(X,Y2), turn_left_into(X,Y3), drive_straight_ahead(X)).

```

In der in dem Programm enthaltene Anfrage zur Quantorbestimmung werden folgende Mengen definiert: Die Menge aller Fahrzeuge, die sich irgendwann im Laufe des Beobachtungszeitraums auf der betrachteten Kreuzung befinden (*on\_intersection(X)*); die Menge aller Fahrzeuge, die irgendwann aus der Straße *Y1* kommend auf die Kreuzung fahren (*come\_from(X, Y1)*), sowie die Mengen aller in bestimmte Straßen rechtsabbiegenden, linksabbiegenden bzw. geradeausfahrenden Fahrzeuge (Prädikatsnamen entsprechend). Somit unterscheiden sich beide Regeln nur in ihrem Kopf, wobei beim Prädikat *come\_from(X, Y1)* im Gegensatz zum Prädikat *on\_intersection(X)* die Bezugsfahrbahn mit erfaßt wird. Die Vorgehensweise, zur Quantorbestimmung auf einen Schlag den gesamten Beobachtungszeitraum zu berücksichtigen (durch die Verwendung des existentiellen temporalen Operators *event*), ermöglicht eine existentielle Quantorberechnung. Bei einer zeitpunktweise fortschreitenden Berechnung wäre dies nicht möglich. Mit Hilfe des Prädikats '*quantorbestimmung*' werden folgende Mengenverhältnisse untersucht: Der Anteil derjenigen Fahrzeuge auf der Kreuzung, die aus der Straße *Y1* kommen; der Anteil derjenigen aus *Y1* kommenden Fahrzeuge, die nach rechts in die Straße *Y2* fahren; der Anteil derjenigen aus *Y1* kommenden Fahrzeuge, die nach links in die Straße *Y3* fahren, und der Anteil derjenigen aus *Y1* kommenden Fahrzeuge, die geradeaus fahren. Dabei wird in den Fällen, in denen die Gesamtzahl der anteiligen Fahrzeuge die Zahl '10' übersteigt, mit Hilfe des Prädikats '*quantorisierung*' dem entsprechenden Zahlenverhältnis der zugehörige Quantor aus Abbildung 3.7 zugeordnet.

### 3.5.2.2 Zahlwörter

Eine andere Art der Quantorisierung, also der Zusammenfassung vergleichbarer Geschehen zu einer quantorisierten Aussage, ist die Verwendung von Zahlwörtern. Das im vorigen Abschnitt zuletzt angegebene metrisch temporale Logikprogramm erlaubt auch die Assoziation von Zahlwörtern. So wird durch das Prädikat *'quantorisierung'* bei einer anteiligen Fahrzeuganzahl kleiner zehn ein Zahlwortquantor zugeordnet und nur im gegenteiligen Fall ein unscharfer Quantor. Die Berechnung von Zahlwortquantoren ist im Vergleich zur Berechnung unscharfer Quantoren weniger problematisch, da die jeweilige absolute Fahrzeuganzahl den Zahlwortquantor eindeutig bestimmen läßt. Bei großen Fahrzeuganzahlen wird die Verwendung von unscharfen Quantoren generell bevorzugt. So dürfte es für die meisten Anwendungen nebensächlich sein, ob 1.442 oder 1.443 Fahrzeuge einen bestimmten Ablauf durchführten. Stattdessen dürfte oft die Feststellung ausreichen, daß es 'sehr viele' waren. Bei kleinen Zahlenverhältnissen ist die Verwendung von Zahlwortquantoren eindeutig vorzuziehen. Bei einem Zahlenverhältnis 'zwei von dreien' beispielsweise ist es unzweckmässig, von 'den meisten' oder 'fast allen' Fahrzeugen zu sprechen. Statt dessen wird man im allgemeinen das Zahlenverhältnis 'zwei von dreien' unter Verwendung dieser Zahlwörter exakt angeben. Die Schwelle, ab wann unscharfe Quantoren und nicht mehr Zahlwörter verwendet werden, wurde in dem zuletzt angegebenen Programm des vorigen Abschnitts vorläufig auf zehn anteilige Fahrzeuge gesetzt. Die Entscheidungsfindung, in welchen Fällen welche Art von Quantoren verwendet werden, bedarf aber sicher noch einiger Untersuchungen. Im folgenden ist abschließend die Definition der Zahlwörter von Null bis Zehn in metrisch temporaler Logik angegeben.

```
always (zahlenname(Card,no) :- Card = 0.0).
always (zahlenname(Card,one) :- Card = 1.0).
always (zahlenname(Card,two) :- Card = 2.0).
always (zahlenname(Card,three) :- Card = 3.0).
always (zahlenname(Card,four) :- Card = 4.0).
always (zahlenname(Card,five) :- Card = 5.0).
always (zahlenname(Card,six) :- Card = 6.0).
always (zahlenname(Card,seven) :- Card = 7.0).
always (zahlenname(Card,eight) :- Card = 8.0).
always (zahlenname(Card,nine) :- Card = 9.0).
always (zahlenname(Card,ten) :- Card = 10.0).
```

# Kapitel 4

## Natürlichsprachliche Schicht

Resultat der signalnahen und begrifflichen Auswertung ist eine Menge temporallogischer Fakten, welche die in der untersuchten Szene erkannten Abläufe beschreibt. Aufgabe der natürlichsprachlichen Schicht ist es in erster Linie, aus dieser begrifflichen Beschreibung einen natürlichsprachlichen Text zu erzeugen. Der Inhalt dieses Textes wird durch die Semantik der Fakten bereits vorgegeben. Die Entwicklung der natürlichsprachlichen Schicht berührt somit hauptsächlich linguistische Problemstellungen. Zwei Problemstellungen sollen an dieser Stelle kurz angedeutet werden: Ziel ist es erstens, einen leicht verständlichen Text zu erstellen, wobei zweitens gewährleistet werden soll, daß bei der Erzeugung des Textes die Semantik der zugrundeliegenden Faktenmenge sinngetreu in den Text übertragen wird.

### 4.1 Literaturüberblick

[Dance et al. 96] widmen sich, wie die vorliegende Arbeit, der Auswertung von Realweltbildfolgen von Straßenverkehrsszenen und der Formulierung der Auswertungsergebnisse als natürlichsprachlichen Text. Das System sieht jedoch keine Trennung zwischen natürlichsprachlicher Schicht und Schlußfolgerungsmechanismus vor. Stattdessen wird die Ausprägung eines bestimmten Knotens des zur Ablaufmodellierung verwendeten Netzwerks in einen natürlichsprachlichen Satz umformuliert. Auf diese Weise entstehen als ‘natürlichsprachliche Beschreibung’ eine Abfolge stereotyper Sätze, die eine lose Aufzählung beobachteter Ereignisse darstellt, wobei der entstehende Text keinen inneren Zusammenhang aufweist. Der auf diese Weise generierte Text ist dadurch nur unwesentlich anschaulicher als eine simple Aufzählung begrifflicher Beschreibungen. Auf diese Weise kann von diesem System nur eine einzige Sichtweise einer Szene berechnet werden.

[Howarth&Buxton 98] unterscheiden zwischen der *Beschreibung* und der *Beobachtung* eines Ablaufs. Die *Beschreibung* resultiert in einer unstrukturierten Aufzählung aktuell extrahierter Abläufe. Die *Beobachtung* einer Szene bezweckt das gezielte Aufspüren be-

stimmter, vom Benutzer vordefinierter Abläufe. Mit der Unterscheidung von Beschreibung und Beobachtung tangieren die Autoren Fragestellungen, die in der vorliegenden Arbeit unter dem Begriff 'Perspektivität' betrachtet werden, ohne daß sie jedoch dazu bisher eine breite Grundlage anstreben. Die Beschreibungen und Beobachtungen werden durch natürlichsprachliche Begriffe dargestellt.

Die Selektions- und die Sprachgenerierungskomponente des im Rahmen des Projekts VITRA ([Herzog 92]) entwickelten Systems SOCCER ([André et al. 88]) nimmt Aufgaben wahr, die in unserem System der natürlichsprachlichen Schicht zugeordnet werden können. Allerdings unterscheidet sich auch bei SOCCER die Zielsetzung, aus der natürlichsprachliche Beschreibungen erzielt werden können, von derjenigen des vorliegenden Ansatzes. Während wir unterschiedliche Sichtweisen einer Szene berechnen wollen, beschränkt sich SOCCER auf eine einzige Sichtweise, wobei eine inkrementelle Beschreibung berechnet wird, die stilistisch an die Kommentare eines Radioreporters erinnert, also eines nicht in den Ablauf verwickelten Beobachters, dessen Aufgabe es ist, schritthaltend mit dem beobachtbaren Ablauf eine Beschreibung zu liefern. Diese Beschreibung soll ihrem Adressaten, der den Ablauf selbst nicht sieht, ein möglichst gutes 'Bild' der Szene vermitteln.

Die *Ereigniserkennungskomponente* von SOCCER assoziiert zu den vorliegenden geometrischen Daten Ereignisbeschreibungen. Die *Selektionskomponente* wählt diejenigen Ereignisse aus, die verbalisiert werden sollen, und führt diese der *Sprachgenerierungskomponente* zu (vgl. [Herzog&Wazinski 94], [Herzog&Rohr 95]). Bei der Formulierung eines Satzes aus mehreren Ereignispropositionen erfolgt zuerst die Auswahl eines geeigneten Verbs aus dem zugrundeliegenden Begriffslexikon und die Festlegung der Verbform. Hierzu erfolgt ein Rückgriff auf die Selektionskomponente, die entscheidet, welche Information zur korrekten Identifikation der Bezugsobjekte verbalisiert werden soll. Desweiteren wird der Ereigniszeitraum durch die Verbzeit und die Verwendung von Zeitadverbien angezeigt. Räumliche Information wird durch die Verwendung räumlicher Präpositionen kodiert. Interne Objektidentifikatoren werden in Nominalphrasen übertragen. Ergänzend werden zusätzliche Attribute ausgewählt, welche dem Empfänger die eindeutige Identifikation des betroffenen Objekts ermöglichen sollen. Dies erfolgt wiederum durch die Verwendung räumlicher Relationen (z.B. 'Das linke Tor'), oder durch das nochmalige Erwähnen von bereits vergangenen Ereignissen (z.B. 'Der attackierte Spieler'). Die Kriterien, die der Entscheidung für ein Verb oder ein Temporaladverb zugrunde liegen, werden nicht detailliert beschrieben und scheinen heuristisch motiviert zu sein. Für die Texterzeugung wird auf ein nicht näher ausgeführtes Schablonenbasiertes Verfahren ('Templates') zurückgegriffen.

Durch Ausnutzung der Ergebnisse des Subsystems REPLAI [Retz-Schmidt 92] können auch Intentionen bestimmter Agenten formuliert werden. Dies erfolgt bisher jedoch nur durch explizites Formulieren der Intention (z.B. 'Sie will A tun'). Die Zukunft als Zeitform (z.B. 'Sie wird A tun') wird bisher ebensowenig verwandt wie die Benutzung von Verben, die eine Absicht implizieren (z.B. 'verfolgen'). Allerdings können bereits Beziehungen zwischen Intentionen und tatsächlich beobachteten Handlungen hergestellt



werden (z.B. ‘Sie tat B, um A zu erreichen’).

Das Subsystem ANTLIMA ([Schirra&Stopp 93], [Schirra 94], [Stopp & Blocher 97]) des Systems SOCCER verkörpert das Hörermodell, d.h. in ANTLIMA wird die Fragestellung untersucht, ob ein Benutzer, der auf die von SOCCER generierten natürlichsprachlichen Beschreibungen von Fußballszenen zugreift, diese im beabsichtigten Sinne verstehen wird, oder ob die Beschreibungen Mißinterpretationsmöglichkeiten oder Inkonsistenzen enthalten. Treten signifikante Unterschiede zwischen dem tatsächlichen und dem von ANTLIMA antizipierten Sachverhalt auf, so werden diese Unterschiede von ANTLIMA in geeigneter Weise durch Durchlaufen einer Antizipationsrückkopplungsschleife ([Stopp 92]) korrigiert. Die Rücktransformation einer durch Bildfolgenauswertung extrahierten begrifflichen Beschreibung in eine synthetische Bildfolge wird im Rahmen der vorliegenden Arbeit in [Jeyakumar 98] und [Nagel et al. 99] untersucht.

Das Dialogsystem PRACMA ([Jameson et al. 94], [Ndiaye 96]) simuliert den ‘unkooperativen’, d.h. in der jeweiligen Zielsetzung divergierenden, Dialog eines Käufers mit einem Verkäufer beim Gebrauchtwagenkauf. Das System kann wahlweise eine der beiden Rollen übernehmen. Es überträgt deutschsprachige Sätze in eine semantische Repräsentationsstruktur, auf welche das pragmatische Dialoggedächtnis zugreift. Der Dialogplaner stellt aufgrund des bestehenden Kontexts verschiedene Dialogpläne zusammen, die Vorgaben über den weiteren Verlauf des Dialogs enthalten, abhängig von der Rolle des Systems im Dialog.

[King 96] versucht, bestehende sprachverarbeitende Systeme hinsichtlich der Qualität der von ihnen generierten natürlichsprachlichen Texte nach eigenen Bewertungsgrundsätzen zu bewerten, gesteht jedoch ein, daß eine faire Beurteilung solcher Systeme durch vereinheitlichte Tests selten möglich ist, da sich Arbeitsweisen, Ziele und Arbeitsschwerpunkte dieser Systeme oft grundsätzlich unterscheiden.

[Noy&Hafner 97] entwickeln ein Rahmenwerk, anhand dessen verschiedene Projekte, die sich mit dem Aufbau von Ontologien beschäftigen, verglichen werden können. Diesen Vergleich führen die Autoren anhand zehn bestehender Ontologien durch. Wie auch [King 96] stellen diese Autoren fest, daß ein solcher Vergleich sich bisher nur sehr schwer ziehen läßt. Es bestehen bisher fundamentale Unterschiede bei der Konstruktion von Ontologien und in der Sicht der Welt.

Oben zitierte Arbeiten, die sich mit der Versprachlichung ihrer sensorischen Wahrnehmungen beschäftigen, berücksichtigen den Begriff der Perspektive nur unzureichend. Meist geben sie alle Beobachtungen, die sie treffen, an den Benutzer weiter, der dann abermals vor dem Problem steht, sich seine eigene Übersicht über das extrahierte Datenmaterial zu verschaffen. Andere Systeme strukturieren und abstrahieren gezielt, um eine übersichtliche Darstellung bereitzustellen. Sie beschränken sich dabei jedoch implizit auf eine Perspektive. Es werden implizite Annahmen getroffen, was beschreibungsrelevant ist, darüberhinausgehende Beschreibungen werden nicht ermöglicht.

Die vorliegende Arbeit will die Generierung natürlichsprachlicher Szenenbeschreibungen ermöglichen, die spezifischen Benutzerwünschen angepaßt sind. Dabei spielen die

bereits eingeführten Begriffe der Standpunktbestimmung und der Perspektive eine zentrale Rolle. Auswärtige Arbeiten über die generische Modellierung von Perspektivität bei der Generierung natürlichsprachlicher Beschreibungen finden sich in der Literatur dementsprechend nicht. Eine Ausnahme bildet [Leong 98], welcher ein System zur Entscheidungsfindung entwickelt, das seine Entscheidungen dadurch trifft, daß es den beobachteten Vorgang aus unterschiedlichen Perspektiven betrachtet. Das System wird auf den Diskursbereich Entscheidungsfindung in der Medizin angewendet. Es wird eine Sprache zur Modellierung und Lösung von dynamischen Entscheidungsproblemen entwickelt, die allerdings bisher nicht auf einen Bildauswertungsprozeß aufsetzt.

## 4.2 Linguistische Grundlagen

Um dem Leser, der vermutlich eher mit Problemstellungen der Informatik, speziell der Bildauswertung, vertraut sein dürfte, einen Einblick zu geben, welche Probleme sich bei der Entwicklung einer natürlichsprachlichen Schicht ergeben, folgt ein kurzer Überblick über allgemeine linguistische Fragestellungen. Da ein solch umfassender Forschungsbereich wie die Linguistik nicht in wenigen Zeilen erschöpfend behandelt werden kann, wird der folgende Überblick bewußt kurz gehalten und beschränkt sich auf für diese Arbeit relevante Problembereiche. [Grewendorf et al. 91] liefern eine interessante, ergiebige und einleuchtende Einführung in linguistische Fragestellungen. Aus diesem Werk sind die folgenden Ausführungen in der Hauptsache entnommen.

Als die *vier zentralen Fragen der Sprachwissenschaft* werden von [Grewendorf et al. 91] folgende Fragestellungen bezeichnet:

1. Was heißt es, sprachliches Wissen zu besitzen (im Gegensatz zu den Ausdrucksformen von Tieren)?
2. Was heißt es, ein ganz spezielles sprachliches Wissen zu besitzen (z.B. Kenntnis des Deutschen im Gegensatz zum Englischen o. ä.)?
3. Wie wird sprachliches Wissen erworben?
4. Wie wird das erworbene sprachliche Wissen verwendet?

Beim Entwurf einer natürlichsprachlichen Systemschicht innerhalb eines Bildauswertungssystems, welche die Ergebnisse der Auswertung sensorischer Daten als natürlichsprachlichen Text formulieren soll, sind für uns unter den oben aufgeführten besonders die dritte und vierte Fragestellung relevant. Wenn wir (zumindest im Ansatz) verstehen lernen, wie der Mensch sprachliches Wissen erlangt und anwendet, so können wir Rückschlüsse ziehen, wie wir im Rahmen unseres Systemansatzes sprachliches Modellwissen repräsentieren und nutzen können. Leider ist die linguistische Forschung in der Beantwortung dieser Fragen noch nicht allzu weit gediehen. Wenn man sich überlegt, wie Kinder ihre Sprache lernen, so ist man geneigt, sich dieses Lernen nach folgendem

Modell vorzustellen ([Grewendorf et al. 91]): Kinder werden mit sprachlichen Daten konfrontiert, leiten daraus Informationen ab und generalisieren diese. Die Reaktion der Umwelt verifiziert oder falsifiziert diese Generalisierungen. Gegen dieses vereinfachte Modell eines *empiristischen* Spracherwerbs spricht (vgl. [Grewendorf et al. 91]):

1. Die sprachlichen Daten, mit denen Kinder konfrontiert werden, sind größtenteils defekt (z.B. aufgrund unzureichender Gedächtnisleistung bei komplizierten Satzkonstruktionen, Versprecher, idiosynkratische Sprachverwendungen) und sind deshalb zu Generalisierungszwecken völlig ungeeignet.
2. Die Daten sind nicht nur defekt, es sind auch viel zu wenige.

Dennoch kann das Kind Sätze bilden, mit denen es noch nie konfrontiert war; es kann noch nie gehörte Worte richtig aussprechen; es kann am Klang eines ihm unbekanntes Wortes erkennen, ob es sich um ein deutsches Wort handeln kann oder nicht; es kann Pronomen den richtigen Subjekten und Objekten eines Satzes zuordnen, selbst wenn es diesen Satz zuvor noch nie gehört hat. Die empiristische Theorie kann nicht nur nicht erklären, warum Sprecher einer Sprache immer neue Sätze korrekt bilden können, sie kann zudem auch nicht erklären, warum bestimmte offensichtliche Generalisierungen nicht vorgenommen werden. Letztgenanntes Phänomen soll an folgendem englischsprachigen Beispiel illustriert werden (aus [Grewendorf et al. 91]). Man betrachte folgenden Satz: *Peter kissed Mary*. Um das Objekt dieses Satzes zu ermitteln, kann man folgende Frage formulieren: *Who did you say that Peter kissed?* Den gleichen Satzbau könnte man jetzt verwenden, um das Subjekt des Satzes zu erfragen. Es entsteht dadurch die Formulierung: *Who did you say that kissed Mary?* Diese Frage ist ungrammatisch. Es gibt im empiristischen Sprachmodell weder eine Erklärung dafür, warum diese Generalisierung blockiert sein sollte, noch ist es vorstellbar, daß der Sprachlerner all das relevante Training erhalten haben soll, um die Generalisierung auf ungrammatische Formulierungen verhindern zu können. Eine Alternative zu dem wenig geeigneten empiristischen Spracherwerbsmodell stellt das *natavistische* Modell dar. Es besagt im wesentlichen, daß der Mensch von Natur aus mit einem ‘geistigen Apparat der Sprachfähigkeit’ ausgestattet ist, dessen ‘Mechanismus durch Konfrontation mit Erfahrungsdaten lediglich ausgelöst’ wird. In Konsequenz wird die Sprache dabei als ein *Organ* des menschlichen Organismus angesehen, Sprache wird folglich nicht erlernt, sondern erworben. Des Menschen Sprachfähigkeit ist dabei weitaus größer als der alltägliche Sprachgebrauch zeigt. Zum Beispiel ist der Mensch in der Lage, folgenden Satz zu bilden:

*Derjenige, welcher denjenigen, welcher den Pfahl, welcher an der Brücke, welche über den Fluß führt, steht, umgeworfen hat, anzeigt, erhält eine Belohnung.*

Dieser Satz ist grammatisch korrekt und auch inhaltlich verstehbar. Trotzdem wird ein solcher Satz im Alltag aus naheliegenden Gründen nicht verwendet. Die Fähigkeit, solche Sätze bilden zu können, wird als *sprachliche Kompetenz* bezeichnet. Das Ausmaß, wie weit wir von dieser Fähigkeit Gebrauch machen (bedingt durch Fakto-

ren wie Gedächtnis, Konzentration usw.), wird als *sprachliche Performanz* bezeichnet. Die sprachliche Kompetenz kann dabei unterteilt werden in eine *grammatische Kompetenz* (die Fähigkeit, komplizierte Sätze bilden zu können) und in eine *pragmatische Kompetenz* (die Fähigkeit, Sätze angemessen zu gebrauchen).

Ein zentraler Begriff der Linguistik ist die *Grammatik*. Der Sprecher einer Sprache ist gemeinhin in der Lage, Sätze dieser Sprache weitgehend korrekt zu bilden und Urteile über die grammatische Korrektheit und die Struktur von Sätzen dieser Sprache abzugeben. Eine Grammatik ist ein Modell dieser Kompetenz des Sprechers. Über die Frage, welche Form eine solche Grammatik haben muß, herrscht unter Linguisten keine generelle Übereinstimmung. Weitestgehend unbestritten ist jedoch, daß eine Grammatik folgende Komponenten besitzen muß:

- Im *Lexikon* muß jedes Wort der Sprache aufgeführt sein, zusammen mit Angaben über dessen Bedeutung, Aussprache, interner Struktur, Wortart und seiner Rolle bei der Bildung von Sätzen.
- Die *phonologische Komponente* beinhaltet das Inventar an Lauten und beschreibt, wie Laute zu Wörtern kombiniert werden.
- Die *syntaktische Komponente* beschreibt, wie Wörter zu Wortgruppen und Sätzen kombiniert werden können.
- Die *semantische Komponente* ermöglicht die Ermittlung der Bedeutung eines Satzes aus den Bedeutungen seiner Wörter und deren Konstellation.

[Guthrie et al. 96] geben einen Überblick über verschiedene Ansätze, die sich mit der Frage auseinandersetzen, wie traditionelle Wörterbücher für die maschinelle Verarbeitung natürlicher Sprache genutzt werden können. [Kamp&Rossdeutscher 94a] untersuchen Fragestellungen, die Form und Inhalt eines Lexikons betreffen, welches die systematische rechnerinterne Repräsentation von Texten unterstützen soll.

Die *Phonetik* untersucht, vereinfacht ausgedrückt, wie Sprachen von verschiedenen Lauten Gebrauch machen, so daß letztere Gegenstand der phonetischen Beschreibung sind. Laute sind dann Gegenstand der phonetischen Beschreibung, wenn sie in einer entsprechenden Sprache Bedeutungen ausdrücken. Die *Phonologie* beschäftigt sich mit Aufbau und Funktion der phonologischen Komponente einer Grammatik.

[Grewendorf et al. 91] bezeichnen die adäquate Beschreibung des Strukturaufbaus von Sätzen als Gegenstand der *Syntax*. Angesichts der Komplexität von Sprachen kann dieses Ziel nur in einem bescheidenen Rahmen erreicht werden. Die Struktur von Sätzen liegt nicht unmittelbar auf der Hand, so daß sie den Sätzen selbst direkt anzusehen wäre. Deshalb können syntaktische Beschreibungen je nach Zweck und Methode sehr unterschiedlich ausfallen. Allgemein gilt jedoch, daß bestimmte Aspekte der natürlichen Sprache (wie die Syntax und auch die Semantik) mit denselben formalen Mitteln beschrieben werden können, wie sie in der Mathematik, Logik oder Informatik verwendet

werden. Die *Theorie der formalen Sprachen* besagt, daß demzufolge eine Sprache nichts weiter ist als die Menge von Sätzen, die von einer gegebenen Grammatik erzeugt werden. Ein *Satz* ist dabei eine Kette terminaler Elemente, die, ausgehend vom Symbol *S*, durch Anwendungen der Regeln der Grammatik erzeugt werden kann. *Phrasenstrukturgrammatiken* sind eine Form von Grammatiken zur Erzeugung einer Sprache. Sie bestehen aus *Phrasenstrukturregeln* und *Worteinsetzungsregeln*. Charakteristisch für Phrasenstrukturgrammatiken ist die Erzeugung von Strukturbäumen. Vertiefende Erkenntnisse zu diesen Grammatiken entnehme man [Grewendorf et al. 91] oder [Kamp&Reyle 93]. Andere Verfahren verwenden *Unifikationsgrammatiken* und deren Weiterentwicklungen ([Burkert 96]).

Die *Morphologie* beschäftigt sich mit der inneren Struktur der Wörter einer Sprache, also mit den universellen und sprachspezifischen Regularitäten, die auf einer Strukturebene zwischen Phonologie und Syntax liegen. Der zentrale Begriff *Morphem* bezeichnet eine *kleinste, in ihren verschiedenen Vorkommen als formal einheitlich identifizierbare Folge von Segmenten, der (wenigstens) eine als einheitlich identifizierbare außerphonologische Eigenschaft zugeordnet ist* (vgl. [Grewendorf et al. 91]).

Die *Semantik* ist dasjenige Teilgebiet der Linguistik, das die Bedeutung sprachlicher Ausdrücke zum Gegenstand hat. In der *Satzsemantik* setzt man voraus, daß es einen engen Zusammenhang zwischen der Bedeutung von Aussagesätzen und deren Wahrheitsgrad gibt. Es gelten vier satzsemantische Prinzipien (vgl. [Grewendorf et al. 91]):

1. Wenn A und B Sätze sind, und A ist wahr und B ist falsch, dann bedeuten A und B nicht dasselbe.
2. Kennt eine Person die notwendige und hinreichende Bedingungen für die Wahrheit bzw. Falschheit eines Satzes, dann kennt diese Person auch die Bedeutung des betreffenden Satzes.
3. Kennt eine Person die Bedeutung eines Satzes, dann sind dieser Person auch notwendige und hinreichende Bedingungen für die Wahrheit bzw. Falschheit des Satzes bekannt.
4. Für jeden Aussagesatz bestimmt die Semantik die Wahrheitsbedingungen. Die Bedeutung eines Satzes (= seine Wahrheitsbedingungen) läßt sich aus den Bedeutungen seiner Teilausdrücke ermitteln.

Eine Semantik, die diese Prinzipien erfüllt, wird *modelltheoretische Semantik* genannt. Zur Bestimmung der Bedeutung werden mathematische Methoden wie die Aussagenlogik und die Prädikatenlogik verwendet.

Als *pragmatische Kompetenz* wurde in (unscharfer) Abgrenzung zur grammatischen Kompetenz die Fähigkeit zur kompetenten Verwendung korrekt gebildeter Sätze in angemessenen Kontexten bezeichnet. Die linguistische Forschung stellt bisher lediglich Ansätze zur Verfügung, um die Frage beantworten zu können, welche kognitiven

Module die pragmatische Kompetenz konstituieren. Dies ist deshalb schwierig, weil dabei soziale, weltanschauliche und andere komplexe Faktoren eine Rolle spielen. Die Vielfalt pragmatischer Fragestellungen und Probleme hat dazu geführt, daß die unterschiedlichsten Vorstellungen und Konzeptionen von Pragmatik ausgebildet wurden. Die verschiedenen Theorien werden in [Grewendorf et al. 91] vorgestellt, aufgrund des Umfangs dieses Materials wird an dieser Stelle jedoch nicht weiter darauf eingegangen. [Akman&Surav 96] vergleichen verschiedene Ansätze zur Formalisierung von Kontext.

## 4.3 Diskursrepräsentationstheorie (DRT)

[Kamp&Reyle 93] widmen sich der Diskursrepräsentationstheorie (DRT), einer wissenschaftlichen Disziplin, die sich mit den Gemeinsamkeiten der Semantik natürlichsprachlicher Äußerungen und Logik auseinandersetzt. Das Hauptanliegen der Autoren ist die Bearbeitung der Frage, wie geschriebene, englischsprachliche Satzfolgen mit Hilfe einer Logikstruktur semantisch repräsentiert werden können. Als solche Logikstruktur entwickeln die Autoren Diskursrepräsentationsstrukturen (DRSen), welche nicht nur einzelne Sätze, sondern auch größere linguistische Einheiten wie Passagen, Abschnitte und Texte repräsentieren können. Durch die Verwendung der DRT in dieser Arbeit ergibt sich die Möglichkeit, Auswertungsergebnisse unter linguistischen Gesichtspunkten rechnerintern zu repräsentieren. Als formal ausgereifter Ansatz bietet die DRT eine geeignete Schnittstelle zwischen den Auswertungsstufen, die die klassische Bildfolgenauswertung betreffen, und den linguistischen Anforderungen, die sich in der natürlichsprachlichen Auswertung ergeben. Desweiteren wird sich zeigen, daß die DRT nach [Kamp&Reyle 93] als Grundlage der natürlichsprachlichen Schicht eine methodische Parallele zu den Werkzeugen der beiden anderen Schichten aufweist: dem Kalmanfilter der geometrischen Schicht und der Situationsanalyse in der begrifflichen Schicht. Allen dreien ist die inkrementelle Vorgehensweise gemein. Die Kalmanfilterung führt anhand einer gegebenen Zustandsschätzung eine Prädiktion zum nächsten Zeitpunkt durch, die dann durch einen Aktualisierungsschritt angeglichen wird (vgl. Abschnitt 2.7). Die Situationsanalyse verfährt ähnlich: Anhand der aktuell bestimmten Situation wird, basierend auf einer Handlungserwartung, zum nächsten Zeitpunkt eine Nachfolgesituation bestimmt. Und auch bei der DRT ist die Vorgehensweise vergleichbar. Die Einfügung einer neuen Aussage in die DRS ist abhängig von der in der DRS zuvor eingebrachten Information.

### 4.3.1 Definition von DRSen

Eine **DRS** ist ein Paar  $\langle U, Con \rangle$ , bestehend aus einer Menge  $U$  von Diskursreferenten (DRen), welche *Universum* der DRS genannt wird, und einer Menge  $Con$  von DRS-Bedingungen.

Das **Vokabular** von DRSen umfaßt *DRen*, *Namen*, *Prädikate* und *Funktoren*.

- **DRen**

Mit dem Kleinbuchstaben **t** plus positiv ganzzahligem Index werden DRen für *Zeitpunkte und Zeitintervalle* (engl. *times*) bezeichnet.

Mit dem Kleinbuchstaben **e** plus positiv ganzzahligem Index werden DRen für *Ereignisse* (engl. *events*) bezeichnet.

Mit dem Kleinbuchstaben **s** plus positiv ganzzahligem Index werden DRen für *Zustände* (engl. *states*) bezeichnet.

Mit den Kleinbuchstaben **mt** plus positiv ganzzahligem Index werden DRen für *Zeitdauern* (engl. *amounts of time*) bezeichnet.

Alle anderen Kleinbuchstaben, insbesondere **x** und **y** plus positiv ganzzahligem Index, bezeichnen DRen für *Agenten* und *Bezugsobjekte*.

- **Namen** werden mit *PN* plus positiv ganzzahligem Index bezeichnet.

- Als **Prädikate** werden unterschieden:

einstellige Prädikate für Individuen: Großbuchstabe *N* plus positiv ganzzahligem Index;

n-stellige Ereignisprädikate;

n-stellige Zustandsprädikate;

Zweistellige Prädikate für Ereignisse, Zustände und Zeitintervalle;

Einstellige Prädikate für Zeitintervalle: Großbuchstabe *T* plus positiv ganzzahligem Index;

Einstellige Prädikate für Zeitdauern: Großbuchstaben *MT* plus positiv ganzzahligem Index;

- **Funktoren**

Einstellige Funktoren für Ereignisse oder Zustände: *beg*, *end*, *loc*;

Einstellige Funktoren für Ereignisse, Zustände und Zeitintervalle: *dur*.

**Terme** werden von den Autoren folgendermaßen definiert:

- Jeder DR ist ein Term.

- Ist  $\tau$  ein DR für ein Ereignis oder Zustand, dann sind  $beg(\tau)$  und  $end(\tau)$  Terme für Ereignisse und  $loc(\tau)$  ein Term für Zeitintervalle.

- Ist  $\tau$  ein Term für Ereignisse, Zustände oder Zeitintervalle, dann ist  $dur(\tau)$  ein Term für Zeitdauern.

Folgende Ausdrücke werden als DRS-Bedingungen bezeichnet:

- $PN_i(\tau)$  mit  $PN_i$  Name und  $\tau$  Term für Individuen;
- $N_i(\tau)$  mit  $N_i$  Prädikat für Individuen und  $\tau$  Term für Individuen;
- $e_i : \boxed{\Pi(\tau_i \dots \tau_n)}$  mit  $e_i$  DR für Ereignisse,  $\Pi$  n-stelliges Ereignisprädikat und  $\tau_i$  Terme für Individuen;
- $s_i : \boxed{\Pi(\tau_i \dots \tau_n)}$  mit  $s_i$  DR für Zustände,  $\Pi$  n-stelliges Ereignisprädikat und  $\tau_i$  Terme für Individuen;
- $\tau R \sigma$  mit  $\tau, \sigma$  DREN für Ereignisse, Zustände und Zeitintervalle und  $R$  zweistelliges Prädikat für Ereignisse, Zustände und Zeitintervalle;
- $\Pi(\tau)$  mit  $\Pi$  einstelliges Prädikat für Zeitintervalle und  $\tau$  Term für Zeitintervalle;
- $\Pi(\tau)$  mit  $\Pi$  einstelliges Prädikat für Zeitdauern und  $\tau$  Term für Zeitdauern;
- $\tau = \sigma$  mit  $\tau, \sigma$  Terme des gleichen Typs;
- $\neg K, K_1 \Rightarrow K_2, K_1 \begin{array}{c} \diamond \\ Q \\ x \end{array} K_2$  mit  $K, K_1, K_2$  DRSen,  $x$  DR und  $Q$  verallgemeinerter Quantor.

Auf die Bedeutung der einzelnen Bezeichnungen wird an späterer Stelle genau dann eingegangen, wenn sie für unsere Zwecke von Belang sind. Schon allein die in den obigen Definitionen verwendete Begriffswahl legt nahe, daß die DRS-Notation eine Variante der Prädikatenlogik erster Stufe ist. Dabei übertrifft die DRS-Notation die Prädikatenlogik dahingehend an Mächtigkeit, daß eine DRS ein Paar zweier Mengen ist, welche unendlich groß sein können. Die Transformation einer DRS in einen prädikatenlogischen Ausdruck erfordert also die Einführung einer Ordnung für besagte Mengen und die Beschränkung auf den Fall endlicher Mengen. Im wesentlichen erfolgt dann die Transformation durch Konjunktion der einzelnen DRS-Bedingungen und Übertragung der Schreibweise in die Notation der Prädikatenlogik. [van Eijck&Kamp 97] erweitern den Ansatz aus [Kamp&Reyle 93] aufgrund von Anforderungen, die sich durch eine Vergrößerung des bisher berücksichtigten Sprachfragments ergeben. [Reyle 95] erweitert den Formalismus um unterspezifizierte DRSen (UDRSen), die die Behandlung von Mehrdeutigkeiten in sprachlichen Äußerungen unterstützen. [Eberle 1997] umreißt, basierend auf [Kamp&Reyle 93], ein allgemeines System zur Analyse deutschsprachiger Texte, ausgehend von einem breiten Fragment der deutschen Sprache. Hierzu modifiziert und erweitert der Autor die Theorie der unterspezifizierten Diskursrepräsentation (UDRT).



### 4.3.2 Syntaxanalyse

Die Konstruktion von DRSen erfolgt satzweise. Um zu verstehen, welche Information durch den nächsten Satz eines Diskurses zu dem hinzugefügt wird, was schon an Information aus den vorausgegangenen Sätzen extrahiert wurde, muß der Interpret den neuen Satz zu der Informationsstruktur in Beziehung setzen, die er aus den vorausgegangenen Sätzen erhalten hat. Die Interpretation eines Satzes ist folglich abhängig von der syntaktischen Struktur des Satzes selbst und der Struktur, die den Kontext der anderen Sätze repräsentiert. Sei eine Satzfolge  $S_1, S_2 \dots S_n$  gegeben. Sei  $K_j$  diejenige DRS, die die Semantik der Sätze  $S_1$  bis  $S_j$  ( $1 \leq j \leq n$ ) repräsentiert. Die Einbringung eines Satzes  $S_i$  in die DRS  $K_{i-1}$  erfolgt in zwei Schritten. Im ersten Schritt werden die Ergebnisse der syntaktischen Analyse von  $S_i$  zu der Menge an Bedingungen in  $K_{i-1}$  hinzugefügt. Es entsteht dadurch die DRS  $K$ . Im zweiten Schritt werden *reduzible* Bedingungen aus  $K$  durch Anwendung von Konstruktionsregeln eliminiert. Reduzible Bedingungen sind solche, welche durch Anwendung weiterer Regeln umgeformt werden können. Letztendlich entsteht eine DRS  $K_i$ , die ausschließlich irreduzible (terminale) Bedingungen enthält. Analog zu der Vorgehensweise von [Kamp&Reyle 93] wird auch in unserem Ansatz bei der DRS-Konstruktion bisher kein weiteres a-priori-Wissen genutzt, das heißt, die DRS  $K_0$  ist generell leer. Um aus einem gegebenen Satz DRS-Bedingungen zu generieren, welche dessen Semantik repräsentieren, muß im ersten Schritt dessen Syntax analysiert werden. Dies erfolgt anhand einer Phrasenstrukturgrammatik. Der durch die Syntaxanalyse berechnete Phrasenstrukturbaum bildet die Grundlage für die semantische Analyse, wobei durch die Anwendung spezieller Konstruktionsregeln die DRS aufgebaut wird. Das von den Autoren aufgebaute Konstruktionsregelwerk ist nicht umfassend genug, um jeden erdenklichen englischsprachigen Satz semantisch analysieren zu können. Stattdessen wird nur ein kleiner Ausschnitt englischsprachiger Satzkonstruktionen betrachtet.

### 4.3.3 Ereignisse und Zustände

Eine DRS wird als wahr betrachtet, wenn für jeden DRen ihres Universums Individuen gefunden werden können, so daß die Bedingungen, die die DRS enthält, durch die korrespondierenden DRen eingehalten werden. Die Autoren entwerfen ein recht umfassendes und detailliertes Modell zur Verifikation von DRSen. Für unsere Zwecke von unmittelbarer Bedeutung ist dabei das von den Autoren verwendete Zeitmodell. Eine zentrale Stellung in den temporallogischen Betrachtungen von [Kamp&Reyle 93] nimmt der Begriff *Eventuality* ein, der hier mit dem in der Bildfolgenauswertung zentralen Begriff *Ablauf* gleichgesetzt wird. [Kamp&Reyle 93] unterteilen Abläufe in *Ereignisse* und *Zustände*. Ereignisse beschreiben Abläufe, die eine Änderung eines bestehenden Zustandes beschreiben (z.B. *The vehicle stopped*). Umgekehrt beschreibt ein Zustand das Ergebnis eines Ereignisses (z.B. *The vehicle is standing*). Im Vordergrund steht dabei der Begriff *Ereignis*, wobei Zustände vereinfachend als Vor- oder Nachbedingungen von Ereignissen betrachtet werden. Für ihre weiteren Ausführungen setzen [Kamp&Reyle 93]

die Trennbarkeit von Ereignissen und Zuständen voraus. Diese Arbeitshypothese läßt sich jedoch nicht ohne weiteres aufrecht erhalten, wenn ein automatisches System mit der Repräsentation von Abläufen, insbesondere mit der Zuordnung von Abläufen zu Ereignissen oder Zuständen betraut werden soll. Folgende Beispielsätze sollen die damit verbundenen Probleme illustrieren.

- (11.1) *Objekt\_4 biegt von Straße A nach links in Straße B.*
- (11.2) *Es regnet.*
- (11.3) *Es hat geregnet.*
- (11.4) *Es regnete.*
- (11.5) *Neulich besuchte ich meine Oma. Am ersten Tag gingen wir ins Museum. Am zweiten Tag blieben wir zu Hause. Am dritten Tag war ich krank und am vierten Tag fuhr ich wieder heim.*
- (11.6) *Als mich meine Oma besuchte, war ich krank.*
- (11.7) *Objekt\_5 fuhr in die Tankstelle. Währenddessen stand Objekt\_6 auf dem ersten Tankplatz.*
- (11.8) *Objekt\_6 stand in der Phase auf dem Tankplatz, als Objekt\_5 in die Tankstelle fuhr.*

Intuitiv ist klar, daß man (11.1) eher als die Beschreibung eines Ereignisses und (11.2) als Zustandsbeschreibung auffassen würde. Dennoch ist es schwer, deutlich zu machen, was ein Ereignis und was einen Zustand genau ausmacht. Anhand des verwendeten Verbes allein läßt sich die Zuordnung nicht begründen. Zwar dient das Verb ‘*regnen*’ in (11.2) als Teil einer Zustandsbeschreibung, doch kann dieses Verb auch zur Beschreibung eines Ereignisses herangezogen werden, wie z.B. in (11.3). [Kamp&Reyle 93] nennen einige Anzeichen, die die Unterscheidung zwischen Ereignissen und Zuständen erleichtern. Eines davon ist die verwendete Zeitform: Die Verwendung des Perfekts läßt uns (11.3) als die Beschreibung eines vergangenen Ereignisses auffassen, während durch die Verwendung des Imperfekts (11.4) zur Beschreibung eines vergangenen Zustandes wird. Wie schwierig die Trennung zwischen Zuständen und Ereignissen im Einzelfall sein kann, dokumentieren (11.5,6). Beide Diskurse enthalten die Formulierung ‘*war ich krank*’. Doch während dadurch in (11.6) ein Zustand beschrieben wird, handelt es sich in (11.5) dabei um ein Ereignis. Die Zuordnung einer Phrase zu Ereignissen oder Zuständen hängt also auch von kontextuellen Gegebenheiten ab und läßt sich nicht allein aus der Struktur der Phrase ableiten. Die Trennung von Ereignissen und Zuständen ist ein diffiziles und nicht vollständig gelöstes Problem. Dennoch nimmt sie eine wichtige Stellung im Ansatz von [Kamp&Reyle 93] ein, da sie die semantische Repräsentation temporaler Aspekte einer Aussage maßgeblich beeinflußt. Will man die Theorie von [Kamp&Reyle 93] adaptieren, muß man sich folglich dem offenen Problem der Trennung von Zuständen und Ereignissen stellen. Für den im Rahmen der vorliegenden Arbeit untersuchten Ansatz könnte dies bedeuten, daß jedesmal, wenn ein neues Faktum als Teil einer begrifflichen Beschreibung in eine DRS überführt werden

soll, besagte Zuordnungsentscheidung getroffen werden müßte. Im folgenden sind einige Beispiele von extrahierten Fakten angegeben. Da es in diesem Abschnitt letztendlich um die Repräsentation temporaler Aspekte geht, sind die zeitlichen Grenzen für die Gültigkeit dieser Fakten mit aufgeführt.

(11.9) 50 : 100 ! fahren\_in(objekt\_5, die\_Tankstelle, 1).

(11.10) 50 : 100 ! stehen\_auf(objekt\_6, der\_erste\_Tankplatz, 2).

Faktum (11.9) würde man als Beschreibung eines Ereignisses werten, besonders wenn man es als das Äquivalent folgender natürlichsprachlicher Aussage auffaßt: *Objekt\_5 fährt in die Tankstelle*. Faktum (11.10) hingegen erscheint als Zustandsbeschreibung, denn sein analog gebildetes natürlichsprachliches Äquivalent beschreibt einen Zustand: *Objekt\_6 steht auf dem ersten Tankplatz*. Doch der Schein trügt, die Unterscheidung ist so einfach nicht. So enthalten (11.9,10) alle erforderlichen inhaltlichen Informationen, um aus ihnen den Diskurs (11.7) bilden zu können. Hier wird das Einfahren in die Tankstelle wie erwartet als Ereignis und das Stehen auf dem Tankplatz als Zustand geschildert. Doch aus (11.9,10) kann man ebensogut den Diskurs (11.8) bilden, und hier verhält es sich genau umgekehrt: Das Einfahren ist der Zustand und das Stehen auf dem Tankplatz ist das Ereignis. Das Fazit dieses Beispiels lautet: Man kann aus der Formulierung eines Faktums nicht ableiten, ob es einen Zustand oder ein Ereignis beschreibt. Stattdessen kann es sowohl zur Beschreibung eines Zustands als auch der eines Ereignisses herangezogen werden. Hier unterscheidet sich die Interpretation einer begrifflichen Beschreibung von der eines natürlichsprachlichen Textes. Die Zuordnung Zustand/Ereignis ist folglich im Rahmen der vorliegenden Arbeit erst dann sinnvoll, wenn es darum geht, aus der DRS einen natürlichsprachlichen Text abzuleiten. Als Konsequenz werden bei der Erzeugung der DRS Fakten nicht in Ereignisse und Zustände unterschieden, sondern stets als Abläufe betrachtet, wobei Abläufe bereits zu Beginn dieser Arbeit als Sammelbegriff für Ereignisse und Zustände definiert wurden. Wer nun denkt, daß durch die Behandlung von Fakten als Abläufe die Zuordnungsproblematik aus [Kamp&Reyle 93] gelöst oder zumindest elegant umschifft wurde, der sieht sich getäuscht. Die Fakten (11.9,10) enthalten beide temporale Information: Sie wurden beide zwischen dem Zeitpunkt ‘50’ und dem Zeitpunkt ‘100’ beobachtet. Die in ihnen beschriebenen Abläufe traten also ‘gleichzeitig’ auf. Zur Repräsentation temporaler Aspekte verlangen [Kamp&Reyle 93] die Unterteilung von Abläufen in Zustände und Ereignisse. Da wir gerade auf diese Unterteilung verzichtet haben, sind wir zunächst einmal nicht in der Lage, eben geschilderte temporale Aspekte semantisch zu repräsentieren. Doch der Verzicht auf eine semantische Repräsentation temporaler Aspekte würde sowohl temporallogisches Schlußfolgern auf DRS-Ebene als auch als Konsequenz die Konstruktion natürlichsprachlicher Texte in unverletzbarer Weise einschränken. Da aber die Beschreibung temporaler Zusammenhänge zwischen Abläufen weiterhin gewährleistet werden soll, ist die Lösung des Kernproblems unumgänglich, wie man temporale Aspekte von Abläufen repräsentiert, ohne Abläufe in Ereignisse und Zustände unterteilen zu müssen. Hierzu bieten sich drei Lösungsansätze an:

1. Man findet Argumente, die erlauben, temporale Aspekte von Abläufen wie diejenigen von Ereignissen zu behandeln;
2. oder man findet Argumente, aufgrund derer Abläufe sich wie Zustände behandeln lassen;
3. oder man entwickelt eine ganz eigene Theorie zur Behandlung temporaler Aspekte von Abläufen.

Die dritte Alternative wird in dieser Arbeit nicht verfolgt. Die in [Kamp&Reyle 93] entwickelte Theorie wurde als das Produkt vieler Jahre an Erfahrung und linguistischer Forschung gerade deshalb für diese Arbeit ausgewählt, da sie eine fundierte Plattform für die hier verfolgten Zielsetzungen bietet. Es wurde bewußt darauf verzichtet, als Bildauswerter eine eigenen linguistischen Forschungsansatz zu verwirklichen, wie auch einen entsprechenden eigenmächtig zu erweitern. Statt dessen empfiehlt es sich, die Möglichkeiten des gegebenen Ansatzes von [Kamp&Reyle 93] auszuschöpfen. Dies bedeutet, daß die ersten beiden der oben angegebenen Alternativen untersucht werden. Dabei lassen sich tatsächlich Argumente finden, die es erlauben, zum Zwecke der Repräsentation begrifflicher Beschreibungen als DRSen die temporalen Aspekte von Abläufen generell wie diejenigen von Ereignissen zu behandeln (Alternative 1). Man betrachte dazu folgende Beispielsätze:

(11.11) *Es regnet.*

(11.12) *Udo blickte kurz aus dem Fenster. Dann drehte er sich zu seinem Zimmerkollegen um und sagte: "Es regnet."*

(11.13) *Ich war krank.*

(11.14) *Das Fahrzeug will überholen.*

Wie bereits weiter oben ausgeführt, beschreibt (11.11) einen Zustand. Die Äußerung ‘*es regnet*’ in (11.12) hingegen ist die Schilderung eines Ereignisses, das Udo beobachtet hatte, als er kurz aus dem Fenster blickte. Der Grund, daß das Regnen in (11.12) eher Ereignischarakter hat, scheint damit zusammenzuhängen, daß die entsprechende Aussage Ergebnis einer Momentaufnahme ist. Der Aspekt einer gewissen ‘Beständigkeit’, die gewisser Ausdruck jeder Zustandsbeschreibung ist, steht beim Gebrauch der Phrase ‘*es regnet*’ in (11.12) nicht im Mittelpunkt. Ähnlich läßt sich auch für die Sätze (11.13,5,6) argumentieren. Satz (11.13) beschreibt einen Zustand. Auch in (11.6) hat die Phrase ‘*ich war krank*’ Zustandscharakter. Daß in (11.5) dieselbe Phrase plötzlich Ereignischarakter annimmt, scheint damit zu tun zu haben, daß die Phrase Teil einer Aufzählung von Ereignissen ist und dadurch selbst diesen Charakter annimmt. Satz (11.14) beschreibt eine Absicht. Es stellt sich die Frage, ob Absichten Ereignis- oder Zustandscharakter haben. Niemand wird wohl behaupten, daß (11.14) ein Ereignis beschreibt. Denn noch ist nichts passiert, der Überholvorgang soll ja erst noch Wirklichkeit werden und hat sich noch nicht ereignet. Somit könnte man (11.14) als die

Beschreibung eines ‘inneren Zustands’ des Fahrzeug(führer)s beschreiben, der sich in dem Verlangen, überholen zu wollen, äußert. Doch andererseits: Wenn ein Beobachter über ein Fahrzeug, das er beobachtet, die Aussage (11.14) macht, so stellt die Aussage eine Schlußfolgerung aus den beobachteten Ereignissen dar. Der Beobachter hat beispielsweise gesehen, daß das Fahrzeug dichter auffährt (Ereignis), aufblendet (Ereignis) und die Spur wechselt (Ereignis) und zieht daraus die Schlußfolgerung (11.14). Satz (11.14) stellt somit die als Absicht formulierte Zusammenfassung beobachteter Ereignisse dar, und bekommt in dem Falle selbst auch Ereignischarakter. Fazit: Aussagen, die per se starken Zustandscharakter haben, können Ereignischarakter annehmen, wenn sie

- eine *Momentaufnahme* zum Ausdruck bringen (11.12);
- Teil einer Aufzählung von Ereignissen sind (11.5);
- zusammenfassende Schlußfolgerungen von beobachteten Ereignissen darstellen (11.14).

Wie läßt sich diese Feststellung für die semantische Repräsentation von begrifflichen Beschreibungen von Bildauswertungsergebnissen nutzen? Bildfolgenauswertung läßt sich beschreiben als Technik zur Generierung von wissensbasierten Schlußfolgerungen aus mittels bildgebender Sensoren beobachteten raum-zeitlichen Veränderungen in einer Szene. Der zentrale Begriff der ‘Veränderung’ spiegelt sich eher in dem Begriff ‘Ereignis’ als in dem Begriff ‘Zustand’ wieder. Diese Einschätzung setzt sich auch bei der Ableitung natürlichsprachlicher Texte aus Bildfolgenauswertungsergebnissen fort. Wie bereits in einem der vorangegangenen Kapitel festgestellt, besitzt jede natürlichsprachliche Beschreibung einen Erzähler, welcher in erster Linie *Veränderungen*, also Ereignisse, beobachtet und schildert. Somit haben natürlichsprachliche Beschreibungen in erster Linie eine Form, die an (11.5), eine Aufzählung von Ereignissen, erinnert. Zwar können in natürlichsprachlichen Beschreibungen auch Zustandsschilderungen auftreten, diese stellen jedoch oft nur Momentanbetrachtungen dar. Zur Untermauerung dieser These betrachte man noch einmal (11.9,10). Man nehme an, ‘Objekt\_5’ sei dasjenige Fahrzeug, das im Mittelpunkt der Beschreibung stehen soll. Das Verhalten von ‘Objekt\_6’ ist somit nur dann beschreibungsrelevant, wenn es das Verhalten von ‘Objekt\_5’ beeinflußt, also im vorliegenden Beispiel erst dann, wenn ‘Objekt\_5’ sich auf das Tankstellengelände begibt. Daß ‘Objekt\_6’ schon einige Zeit zuvor auf dem Tankplatz gestanden haben dürfte, ist in diesem Fall unwesentlich. Somit bringt Faktum (11.10) nur eine Momentaufnahme mit Ereignischarakter zum Ausdruck. Sei nun umgekehrt ‘Objekt\_6’ Hauptakteur der Beschreibung, und Objekt\_5 Nebenakteur. Bevor ‘Objekt\_6’ besagten Tankplatz eingenommen hat, dürften zuvor andere Ereignisse dieses Fahrzeugs beobachtet worden sein (z.B. das Heranfahen an den Tankplatz). Das Stehen auf dem Tankplatz ist also im Sinne von (11.5) ein Ereignis von mehreren und hat somit Ereignischarakter. Will man explizit zum Ausdruck bringen, daß ‘Objekt\_6’ genau dann auf dem Tankplatz stand, als ‘Objekt\_5’ in die Tankstelle einbog, würde

man angesichts der Tatsache, daß ‘Objekt\_6’ der Hauptakteur ist, eher Formulierung (11.8) als (11.7) wählen. Somit hat das Stehen auf dem Tankplatz wieder Ereignischarakter.

Zusammenfassend lassen sich die vorstehenden Ausführungen in folgenden einfachen Sätzen zusammenfassen: Bei der Erzeugung einer DRS aus einer begrifflichen Szenenbeschreibung wird aus den dargelegten Gründen auf die strikte Trennung von Ereignissen und Zuständen formal verzichtet, indem grundsätzlich der übergeordnete Begriff ‘Ablauf’ verwendet wird. Die weiteren Schlußfolgerungen dieses Abschnitts führten zum Ergebnis, daß die Repräsentation temporaler Informationen, wie [Kamp&Reyle 93] sie für Ereignisse vorsehen, im Rahmen der vorliegenden Arbeit zweckmäßigerweise auf alle Abläufe ausgedehnt wird. Somit läßt sich zum einen der temporale Aspekt eines Ablaufs in einer DRS repräsentieren, zum anderen bleibt die Option offen, bestimmte begriffliche Fakten als Zustandsbeschreibungen zu nutzen.

#### 4.3.4 Zeit

Als Zeitmodell verwenden die Autoren weder ein zeitpunktbasiertes noch ein intervallbasiertes, sondern ein ereignisbasiertes Modell. Der Ereigniszeitraum ist definiert als die zeitliche Ausdehnung eines Ereignisses. Der Ereigniszeitraum muß aufgrund dieser Definition nicht zwangsläufig ein Intervall sein, sondern kann auch als die Vereinigung disjunkter Zeitintervalle aufgefaßt werden, welche Lücken enthalten kann, wie das folgende Beispiel zeigt: *The vehicle stopped*. Der Ereigniszeitraum dieses Ereignisses umfaßt beispielsweise die Zeitspanne von dem Augenblick, wenn der Fahrer erstmals mit dem Schuh das Bremspedal berührt, bis zu dem Augenblick, zu dem das Fahrzeug endgültig steht. Die Gültigkeit des Ereignisses bleibt unverändert, selbst wenn der Fahrer für einen kurzen Augenblick den Fuß von der Bremse nimmt. Das Ereignis fungiert damit als *wahres Argument der (zeitlichen) Aussage*. Grundlage des ereignisbasierten Zeitmodells ist der Begriff *Ereignisstruktur*  $\mathcal{E} = \langle E, <, \circ \rangle$ , wobei  $E$  eine Menge von Ereignissen darstellt, für welche die Relationen ‘<’ (zeitliches Aufeinanderfolgen) und ‘ $\circ$ ’ (zeitliche Überlappung) folgende Forderungen erfüllen:

$$(P_1) \quad e_1 < e_2 \rightarrow \neg e_2 < e_1$$

$$(P_2) \quad e_1 < e_2 \ \& \ e_2 < e_3 \rightarrow e_1 < e_3$$

$$(P_3) \quad e \circ e$$

$$(P_4) \quad e_1 \circ e_2 \rightarrow e_2 \circ e_1$$

$$(P_5) \quad e_1 < e_2 \rightarrow \neg e_2 \circ e_1$$

$$(P_6) \quad e_1 < e_2 \ \& \ e_2 \circ e_3 \ \& \ e_3 < e_4 \rightarrow e_1 < e_4$$

$$(P_7) \quad e_1 < e_2 \vee e_1 \circ e_2 \vee e_2 < e_1$$

Der Begriff *ereignisbasiertes Zeitmodell* besagt nicht, daß Zeitpunkte und Zeitintervalle in den Erwägungen von [Kamp&Reyle 93] keine Rolle spielen, sondern erlaubt die Definition auch dieser Begriffe. Sei  $\mathcal{E}$  Ereignisstruktur. Dann gilt:

1.  $I$  ist ein *Zeitpunkt* aus  $\mathcal{E}$ , wenn gilt:
 
$$I \subseteq E;$$

$$e_1, e_2 \in I \Rightarrow e_1 \circ e_2,$$

$$H \subseteq E, I \subseteq H, \forall e_1, e_2 \in H : e_1 \circ e_2 \Rightarrow H \subseteq I;$$
2. Ein Ereignis  $e$  *ereignet sich zu einem Zeitpunkt*  $I$ , wenn  $e \in I$ .
3.  $\exists e_1 \in I_1, \exists e_2 \in I_2 : e_1 < e_2 \Rightarrow I_1 <_I I_2$ .

Die Relation  $<_I$  impliziert eine lineare Ordnung der Zeitpunkte. Auf diese Weise entsteht eine *Zeitpunktstruktur*  $\mathcal{I}(\mathcal{E}) = \langle I(\mathcal{E}), <_I(\mathcal{E}) \rangle$ .  $\mathcal{Int}(\mathcal{I}) = \langle \mathcal{Int}, <_p, \circ_p \rangle$  heißt Intervallstruktur abgeleitet aus  $\mathcal{I}$ , wenn gilt:

1.  $\mathcal{Int}$  ist Menge konvexer Untermengen von  $I$ , d.h. wenn  $i_1, i_2 \in I$  und  $i_1 <_I i_3 <_I i_2$ , dann gilt:  $i_3 \in I$ .
2.  $X <_p Y$ , wenn für alle  $i_1 \in X$  und  $i_2 \in Y$  gilt:  $i_1 < i_2$ .
3.  $X \circ_p Y$ , wenn  $X \cap Y \neq \{\}$ .
4.  $X \subseteq_p Y$ , wenn für alle  $i \in X$  gilt:  $i \in Y$ .

## 4.4 Einbettung der DRT in Bildfolgenauswertung

Das Hauptanliegen von [Kamp&Reyle 93] liegt darin, den semantischen Gehalt eines natürlichsprachlichen Textes rechnerintern repräsentieren zu können. Eine Systemkomponente, die die von diesen Autoren vorgestellten Methoden realisiert, kann als Teil eines Empfängers betrachtet werden, welcher von einem systemfremden Sender einen Text erhält, dessen Semantik analysiert und diese repräsentiert. In dieser Konstellation werden keine Vorannahmen über den Inhalt des empfangenen Textes getroffen. Daraus folgt, daß bei der Analyse des Textes in besonderer Weise mögliche Mehrdeutigkeiten in der Formulierung erkannt und evaluiert werden müssen. Die Ermittlung der zutreffenden Interpretation einer mehrdeutigen Textpassage kann dabei nur basierend auf der syntaktischen Analyse und der bereits aus dem Text extrahierten Information gelingen. Deshalb widmen [Kamp&Reyle 93] der korrekten Behandlung sprachlicher Mehrdeutigkeiten sehr viel Aufmerksamkeit. Diese Schwerpunktsetzung beeinflusst den Aufbau der von diesen Autoren entwickelten Konstruktionsregeln zur Transformation von Syntaxbäumen in DRS-Bedingungen.

Die Verwendung der Diskursrepräsentationstheorie in unserem Ansatz unterscheidet sich von dieser Modellvorstellung, da die natürlichsprachliche Systemkomponente nicht als Bestandteil eines Empfängers, sondern eines Senders betrachtet wird. Dieser sendet die Ergebnisse seiner sensorischen Wahrnehmungen als natürlichsprachlichen Text an

den Empfänger, welcher in der Regel ein menschlicher Anwender sein wird. Daraus ergibt sich, daß in unserem Ansatz kein systemfremder Text analysiert wird, sondern eine systemintern erstellte Datenmenge. Diese Datenmenge besteht aus einer zeitlich geordneten Folge metrisch temporaler Logikfakten. Es stellt sich die Frage, ob eine solche Menge ebenfalls als Text aufgefaßt werden kann. Dazu muß der Text (also die Gesamtmenge an Fakten) in Sätze zerlegbar sein, zu deren Konstruktion eine Grammatik angegeben werden kann. Dann wäre es möglich, Phrasenstrukturbäume zu berechneten Logikfakten zu generieren. Als nächster Schritt erfolgt analog zu [Kamp&Reyle 93] die Entwicklung von Konstruktionsregeln, welche aus den Phrasenstrukturbäumen DRSEN ableiten. Die Methoden aus [Kamp&Reyle 93] sind folglich auf unseren Forschungsbereich übertragbar. Als Fazit läßt sich feststellen, daß als wesentlicher Unterschied in der von [Kamp&Reyle 93] unterstützten Modellvorstellung systemfremde Texte, und im Gegensatz dazu in unserem Ansatz systemintern erzeugte Texte analysiert und repräsentiert werden. Die Analyse systemintern erzeugter Texte besitzt den Vorteil, daß Mehrdeutigkeiten ausgeschlossen werden können, indem Fakten eindeutig formuliert werden. Für jedes Faktum existiert dabei genau eine Folge von Konstruktionsregeln, die hintereinander ausgeführt die Semantik dieses Faktums extrahiert und in eine DRS überträgt. Die Zahl der bereitzustellenden Konstruktionsregeln wird auf diese Weise stark eingeschränkt und der Aufbau der einzelnen Regeln vereinfacht. [Kamp&Reyle 93] wählen beim Entwurf von Konstruktionsregeln im Rahmen ihrer Arbeit eine phänomenologische Vorgehensweise. Dies bedeutet, daß die Autoren von einfachen Diskursbeispielen ausgehen, für die sie einfache Konstruktionsregeln zur DRS-Erzeugung entwickeln. Im Verlauf ihrer Arbeit werden die Beispiele komplizierter, so daß bis dato entwickelte Konstruktionsregeln revidiert und erweitert werden müssen. Diese Vorgehensweise ermöglicht dem Leser, leicht Einblick in ihm vielleicht nicht so bewußte sprachliche Besonderheiten zu gewinnen, die die DRS-Erstellung maßgeblich beeinflussen. Die gleiche Vorgehensweise wähle ich für meine Arbeit. Es werden einfache Konstruktionsregeln für einfache Beispiele entwickelt, die bei Erhöhung der Komplexität der Beispiele Stück für Stück erweitert werden. Dabei werden, sofern vorhanden, Parallelen zu vergleichbaren Überlegungen aus [Kamp&Reyle 93] aufgezeichnet.

#### 4.4.1 Faktensyntax

Eine von den vorgelagerten Systemschichten generierte begriffliche Beschreibung wird als Text einer Sprache aufgefaßt, wobei jedes Faktum wie ein eigener Satz behandelt wird. Die Sprache, in der begriffliche Beschreibungen formuliert werden, läßt sich durch eine Grammatik angeben. Diese Grammatik ist in Abbildung 4.1 angegeben. Sie wird zur Analyse der Syntax von Fakten herangezogen. Analog zu [Kamp&Reyle 93] werden Strukturbäume erzeugt, anhand derer die Semantik der begrifflichen Beschreibung durch Anwenden von Konstruktionsregeln analysiert wird.

Fakten (F) enthalten üblicherweise zwei- oder dreistellige Relationen. Ein Argument jeder Relation, üblicherweise das letzte, wird als Markierung (MK3) verwendet. Die



F	→	VERBREATION   RAUMRELATION
F	→	QUANTORRELATION   METARELATION
VERBREATION	→	PRAEDNAME ( AG , BO , MK3 ).
VERBREATION	→	PRAEDNAME ( AG , MK3 ).
METARELATION	→	RELNAME ( MK3 , MK1 , MK2 ).
METARELATION	→	RELNAME ( MK3 , MK1 ).
RAUMRELATION	→	RAUMREL ( MK3 , MK1 , BO ).
RAUMRELATION	→	RAUMREL ( AG , BO , MK3 ).
QUANTORRELATION	→	<b>quantor</b> ( ID , GM , UM ).
QUANTORRELATION	→	<b>quantor</b> ( ID , PRAEDNAME ( <b>X</b> , BO ), UM ).
PRAEDNAME	→	V   V PRAEP   AUX V   V ADV   V $V_{sing}$
PRAEDNAME	→	V ADV ADV PRAEP   PRAEP $N_{sing}$   AUX V PRAEP
AG	→	EN   $NP_{sing}$   QP   $NP_{plur}$   PP
$NP_{sing}$	→	ADJ $N_{sing}$   DET $N_{sing}$   DET ADJ $N_{sing}$
$NP_{sing}$	→	DET ORD $N_{sing}$   $N_{sing}$
$NP_{plur}$	→	ADJ $N_{plur}$   DET $N_{plur}$
QP	→	CARD $_{sing}$ $N_{sing}$   CARD $_{plur}$ $N_{plur}$
QP	→	QUANT $_{sing}$ $N_{sing}$   QUANT $_{plur}$ $N_{plur}$
BO	→	EN   $NP_{sing}$   MK1   $NP_{plur}$
MK1	→	MK
MK2	→	MK   ADV
MK3	→	MK   EN
ID	→	QP
GM	→	PRAEDNAME ( <b>X</b> )   PRAEDNAME ( <b>X</b> , BO )
GM	→	PRAEDNAME ( <b>X</b> , <b>Y</b> )
UM	→	PRAEDNAME ( <b>X</b> , BO )
RAUMREL	→	PRAEP
RELNAME	→	EN

Abbildung 4.1: Grammatik zur syntaktischen Analyse von Fakten. Die Bedeutung der Terminal- und Nichtterminalsymbole wird in Tabelle 4.1 erläutert. Die Symbole **X** und **Y** stehen für Variablen, die beim Auswertungsprozeß nicht ausgeprägt und somit in der weiteren Auswertung nicht berücksichtigt werden.

Abkürzung	Bedeutung	Abkürzung	Bedeutung
F	Faktum	AG	Agens
BO	Bezugsobjekt	ID	Individuenbezeichnung
GM	Grundmenge	UM	Untermenge
NP	Nominalphrase	QP	Quantorisierungsphrase
MK	Markierer	V	Verb
PRAEP	Präposition	ADV	Adverb
ADJ	Adjektiv	AUX	Hilfsverb
EN	Eigenname	N	Nomen
DET	Artikel	ORD	Ordinalzahl
QUANT	Quantor	PP	Personalpronomen
CARD	Kardinalzahl	<i>V<sub>ing</sub></i>	Verb in Verlaufsform

Tabelle 4.1: Erläuterung der in Abbildung 4.1 verwendeten Terminal- und Nichtterminalsymbole.

Markierung besteht aus einer natürlichen Zahl als Identifikator, anhand derer das zugehörige Faktum referenziert wird. Die häufigsten Fakten sind diejenigen, die mit einem (Bewegungs-)Verb einen bestimmten Ablauf beschreiben. *Verbrelationen* sind zwei- oder dreistellig. Sie enthalten neben ihrer Markierung auch eine Agensbezeichnung bzw. eine Agensbezeichnung und die Bezeichnung eines zugehörigen Bezugsobjekts (BO, bei transitiven Verben). Räumliche Relationen besitzen statt einer Verbalphrase als Relationsnamen eine räumliche Präposition. Als *Metarelationen* werden Relationen bezeichnet, welche nicht selbst einen Ablauf beschreiben, sondern mehrere Abläufe zueinander in Beziehung setzen. Zur leichteren Unterscheidung bildet bei Metarelationen (siehe (1.3) - (1.5) im folgenden Beispiel) im Gegensatz zu den Verbrelationen jeweils das erste Argument die Markierung, über die das betreffende Faktum referenziert wird. Im folgenden sind einige Beispiele angegeben.

- (1.1) 100 : 200 ! change\_to(object\_4, the\_passing\_lane, 1).
- (1.2) 100 : 200 ! want\_to\_overtake(object\_4, object\_6, 2).
- (1.3) 100 : 200 ! reason(3, 2, 1).
- (1.4) 100 : 180 ! certainty(4, 3, presumably).
- (1.5) 181 : 200 ! confirmation(5, 3).

Die Fakten (1.1, 1.2) enthalten Verbrelationen, deren semantische Bedeutung intuitiv klar sein dürfte. Faktum (1.3) beschreibt einen kausalen Zusammenhang zwischen (1.1) und (1.2). Das Faktum selbst trägt die Markierung ‘3’ und setzt das Faktum mit der Markierung ‘2’ (also (1.2)) mit dem Faktum der Markierung ‘1’ (1.1) in einen kausalen Zusammenhang: Der Wunsch zu überholen (Faktum (1.2)) ist der Grund für den

Spurwechsel in (1.1). Faktum (1.4) liefert ein weiteres Beispiel für eine Metarelation. Man nehme an, die Beschreibung (1.1-3) wurde von einem neutralen Beobachter generiert, der selbst nicht in die beobachteten Abläufe involviert ist. Dieser Beobachter kann die in (1.1) enthaltene Aussage direkt aus seiner Beobachtung ableiten. Da er aber keinen Einblick in das Innenleben der beobachteten Agenten haben dürfte, stellt (1.2) hingegen keine direkte Beobachtung dar, sondern lediglich eine Schlußfolgerung über die Absicht des Agenten, die sich dem Beobachter durch seine eigene Kenntnis der aktuellen Verkehrssituation, durch eigenes Wissen über die übliche Situationsabfolge und durch die Analyse des vorausgegangenen Verhaltens der beobachteten Agenten aufgedrängt hat. Nichtsdestoweniger stellt (1.2) eine mehr oder weniger begründete Vermutung des Beobachters dar. Demzufolge ist auch der unterstellte kausale Zusammenhang zwischen (1.1) und (1.2) in (1.3) lediglich eine Vermutung. Es ist durchaus legitim, in den meisten Anwendungsfällen sogar erwünscht, daß der Beobachter eigene Schlußfolgerungen oder Vermutungen in seine Beschreibung einbringt. Allerdings sollten sie als solche gekennzeichnet sein. Dies geschieht im Beispiel durch Relation (1.4), welche die Markierung ‘4’ enthält und die Aussage von (1.3) als Vermutung kennzeichnet. Solche Vermutungen können im Allgemeinen im weiteren Verlauf bestätigt oder widerlegt werden. Faktum (1.5) repräsentiert die (nachträgliche) Bestätigung der in (1.4) dargelegten Vermutung, nämlich daß es in Kürze zu einem Überholvorgang kommen wird. Eine solche Vermutung wird beispielsweise dadurch bestätigt, daß daraufhin tatsächlich ein solcher Überholvorgang durchgeführt wird. Neben kausalen Beziehungen können noch andere Beziehungen zwischen Teilabläufen hergestellt werden. Ein Beispiel sind Analogiebeziehungen:

- (1.6) 100 : 200 ! be\_occupied(one\_filling\_place, 6).
- (1.7) 100 : 200 ! on(7, 6, the\_upper\_filling\_lane).
- (1.8) 100 : 200 ! be\_occupied(one\_filling\_place, 8).
- (1.9) 100 : 200 ! on(9, 8, the\_lower\_filling\_lane).
- (1.10) 100 : 200 ! analogy(10, 6, 8).

Wieder sind bei Verbrelationen das jeweils letzte Argument und bei allen anderen Relationen das erste Argument Eigenmarkierungen. Die Fakten (1.6-9) sagen aus, daß sowohl auf der unteren als auch auf der oberen Tankspur genau ein Tankplatz besetzt ist. Faktum (1.10) repräsentiert die Analogie der Fakten (1.6,8), die im vorangegangenen Satz mit der Formulierung ‘sowohl ... als auch’ expliziert wurde. Dieses Faktum ermöglicht die Repräsentation einer Ähnlichkeit zwischen Abläufen und die Erwähnung einer solchen in einer darauf aufbauenden natürlichsprachlichen Beschreibung. Neben der Repräsentation von Analogien ist oft auch die Repräsentation ihres Gegenteils, des Nichtvorhandenseins solcher Analogien erwähnenswert:

- (1.11) 100 : 200 ! be\_occupied(one\_filling\_place, 11).
- (1.12) 100 : 200 ! on(12, 11, the\_upper\_filling\_lane).
- (1.13) 100 : 200 ! be\_occupied(two\_filling\_places, 13).
- (1.14) 100 : 200 ! on(14, 13, the\_lower\_filling\_lane).
- (1.15) 100 : 200 ! no\_analogy(15, 11, 13).
- (1.16) *Auf der oberen Tankspur ist ein Tankplatz belegt. Auf der unteren Tankspur sind hingegen zwei Tankplätze belegt.*

Die natürlichsprachliche Beschreibung (1.16) setzt den Belegungszustand der Tankplätze der oberen und unteren Tankspuren zueinander in Bezug. Durch die Vokabel ‘*hingegen*’ wird verdeutlicht, daß sich beide Belegungszustände voneinander unterscheiden. Diese Unterschiedlichkeit wird in der korrespondierenden begrifflichen Beschreibung (1.11-15) durch das Faktum (1.15) repräsentiert. Ein weiteres Beispiel für Beziehungen zwischen Abläufen sind Dekompositionsbeziehungen:

- (1.17) *Object\_4 overtakes object\_6. First, it changes to the passing lane. Then it passes object\_6. Then it changes back to the filling lane.*
- (1.18) 100 : 200 ! overtake(object\_4, object\_6, 18).
- (1.19) 100 : 130 ! change\_to(object\_4, the\_passing\_lane, 19).
- (1.20) 100 : 130 ! first\_part\_of(20, 19, 18).
- (1.21) 131 : 170 ! pass(object\_4, object\_6, 21).
- (1.22) 131 : 170 ! second\_part\_of(22, 21, 18).
- (1.23) 171 : 200 ! change\_back\_to(object\_4, the\_filling\_lane, 23).
- (1.24) 171 : 200 ! last\_part\_of(24, 23, 18).

In (1.17) wird der Überholvorgang unterteilt in drei Teilabläufe. Die Fakten (1.18-23) geben die dazu korrespondierende begriffliche Beschreibung an. Dabei ist beispielsweise Faktum (1.20) wie folgt zu interpretieren: Das Faktum trägt die Markierung ‘20’ und stellt einen Bezug zwischen dem Faktum mit der Markierung ‘19’ und dem Faktum mit der Markierung ‘18’ her. Diese Beziehung besteht darin, daß die in Faktum ‘19’ beschriebene Beobachtung der erste Teil der übergeordneten Beobachtung aus Faktum ‘18’ darstellt.

#### 4.4.2 Eigennamen

[Kamp&Reyle 93] entwickeln anhand einfacher Beispielsätze die Erzeugung von DRSen, die die Semantik des Satzes repräsentieren. Ein solcher Beispielsatz sei: *Jones owns Ulysses*. Dieser Satz wird als zweistellige Relation *own* aufgefaßt, die als Argumente zwei Individuen, *Jones* und *Ulysses*, besitzt. Beide Individuen werden durch ihre Eigennamen bezeichnet, wobei es sich bei erstgenanntem um eine Person, beim zweiten

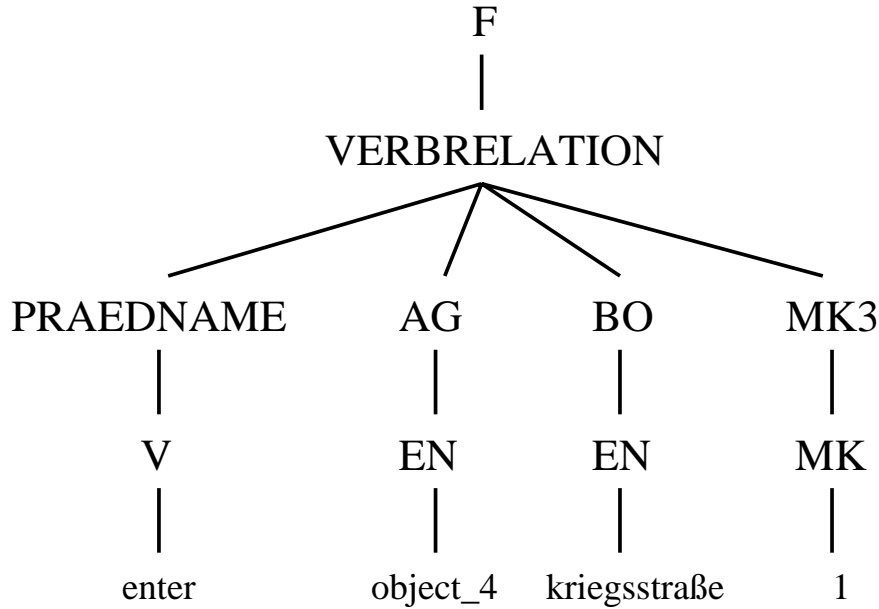


Abbildung 4.2: Strukturbaum zum Faktum **enter(object\_4, kriegsstraße, 1)** nach Anwendung der Konstruktionsregel aus Abbildung 4.3.

um ein Buch handeln dürfte. Zur DRS-Erzeugung werden beide Individuen durch Diskursreferenten (DRen) repräsentiert. Der jeweilige DR und das zugehörige Individuum werden durch eine eigene DRS-Bedingung zusammengeführt. Auch bei der Berechnung begrifflicher Beschreibungen durch Bildfolgenauswertung treten Relationen mit Individuen auf, die durch ihre Eigennamen bezeichnet werden. Man betrachte folgendes Faktum:

$$(2.1) \text{ enter}(\text{object\_4}, \text{kriegsstraße}, 1).$$

Hierbei handelt es sich um eine Relation mit drei Argumenten. Die Angabe des Gültigkeitsintervalls, also der temporalen Komponente der Relation, wird vorerst zur Vereinfachung weggelassen. Die syntaktische Analyse dieser Relation anhand der in Abbildung 4.1 angegebenen Grammatik ergibt den in Abbildung 4.2 aufgeführten Strukturbaum. Bei Faktum (2.1) handelt es sich um eine *Verbrelation*, die ein Verb (*V*) enthält. Als Argumente enthält sie ein Agens (*AG*), ein Bezugsobjekt (*BO*) und eine Markierung (*MK3*). Sowohl das Agens als auch das Bezugsobjekt werden durch Eigennamen (*EN*) bezeichnet. Anhand des Strukturbaums soll die Semantik des Faktums (2.1) in eine DRS übertragen werden. Dazu wird die in Abbildung 4.3 gezeigte Konstruktionsregel zur Behandlung von Eigennamen verwendet. Die Konstruktionsregel gibt an, welche Konfiguration im Strukturbaum vorhanden sein muß, damit diese Regel greift. Ist diese Konfiguration vorhanden, so wird ein neuer DR  $x$  eingeführt. Dieser wird durch eine neue Diskursbedingung  $\alpha(x)$  dem zugehörigen Individuum zugeordnet. Der DR ersetzt dann den angegebenen Teilbaum. Durch Anwenden der Konstruktionsregel KR.EN auf das Faktum (2.1) entsteht demnach die in Abbildung 4.4 gezeigte DRS. Es wird sich

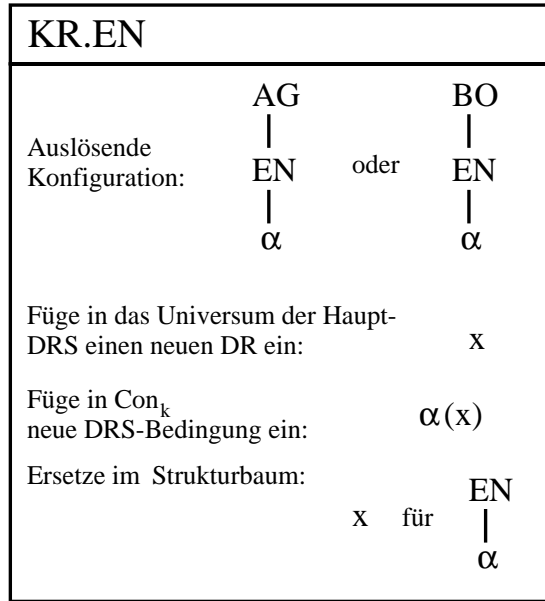


Abbildung 4.3: Konstruktionsregel KR.EN (vorläufige Version).

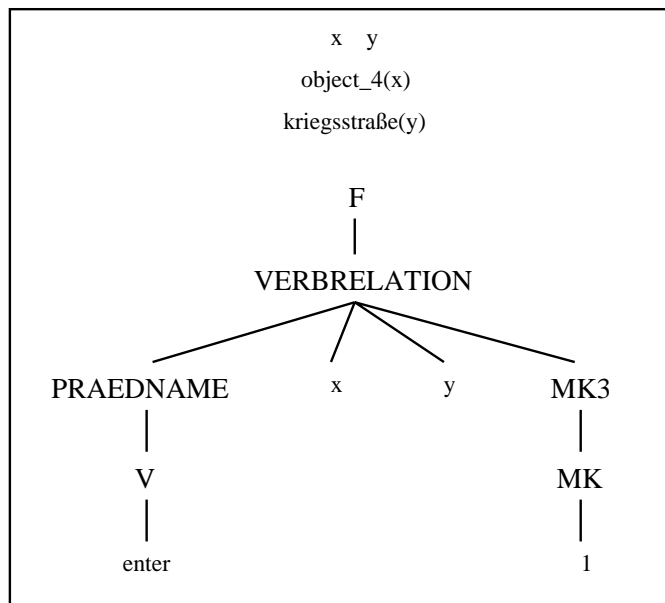


Abbildung 4.4: DRS zum Faktum **enter(object\_4, kriegsstraße, 1)**.

zeigen, daß die Konstruktionsregel KR.EN nicht alle semantischen Schlußfolgerungen berücksichtigt, die bei der Analyse von Fakten nach Art von Beispiel (2.1) getroffen werden können. Um dies zu demonstrieren, soll als nächstes das folgende Faktum in der DRS repräsentiert werden:

(2.2) approach(object\_4, ettlinger\_straße, 2).

Der Strukturbaum zum Faktum (2.2) entspricht im Aufbau dem zum Faktum (2.1). Durch Anwenden der Konstruktionsregel KR.EN entsteht die in Abbildung 4.5 gezeigte DRS. Dabei wird der Ausdruck  $x_1$  *enter*  $y_1$  als verkürzende Schreibweise für den verbliebenen Strukturbaum der DRS aus Abbildung 4.4 verwendet. Diese verkürzende Schreibweise wird, analog zu [Kamp&Reyle 93], im Folgenden stets so verwendet. Es

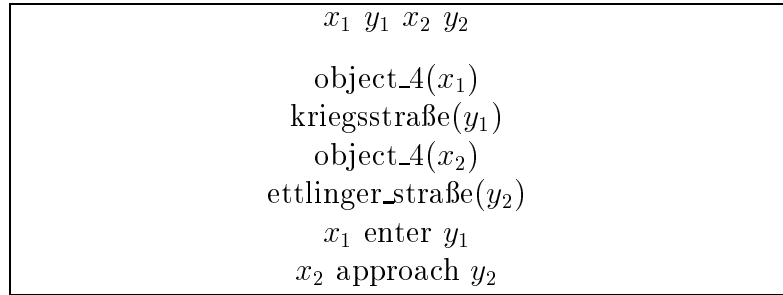


Abbildung 4.5: Diese DRS entsteht durch Hinzufügen des Faktums (2.2) zu der DRS aus Abbildung 4.4. Nähere Erläuterung im Text.

fällt auf, daß die Diskursreferenten  $x_1$  und  $x_2$  dasselbe Individuum bezeichnen. Diese Identität ist jedoch in der DRS nicht unmittelbar dokumentiert. Um dies zu ermöglichen, wird die Konstruktionsregel KR.EN, wie in Abbildung 4.6 gezeigt, erweitert. Falls bereits ein DR  $y$  existiert, der auf das gleiche Individuum  $\alpha$  verweist wie der neue DR  $x$ , dann wird statt der DRS-Bedingung  $\alpha(x)$  die DRS-Bedingung  $x = y$  eingefügt. Durch Anwenden der neuen Konstruktionsregel KR.EN wird aus der DRS aus Abbildung 4.5 die in Abbildung 4.7 gezeigte DRS.

### 4.4.3 Definite Beschreibungen

Definite Beschreibungen sind Individuenbezeichnungen, die mit dem bestimmten Artikel ‘the’ beginnen. Definite Beschreibungen werden generell als eindeutig bestimmbar aufgefaßt. So kann man beispielsweise der definiten Beschreibung ‘*der erste Kaiser des Römischen Weltreiches*’ eindeutig das Individuum ‘*Augustus*’ zuordnen, da generell nur ein Individuum das erste gewesen sein kann, und das war nun einmal besagter Augustus. Nicht immer ergibt sich diese Eindeutigkeit so unmittelbar (vgl. [Kamp&Reyle 93], Seite 248ff.). Man betrachte folgendes Beispiel einer begrifflichen Beschreibung:

$$(3.1) \quad \text{drive\_on}(\text{object\_4}, \text{kriegsstraße}, 1).$$

$$(3.2) \quad \text{be\_fast}(\text{the\_car}, 2).$$

Für sich allein betrachtet ist die Bezeichnung ‘*the\_car*’ aus (3.2) nicht eindeutig, da viele Individuen dieser Bezeichnung genügen können. Durch (3.1) ergibt sich jedoch eindeutig, welches Fahrzeug in (3.2) gemeint ist, nämlich dasjenige, welches den (eindeutigen)

KR.EN		
	AG	BO
Auslösende Konfiguration:	 EN	 EN
	$\alpha$	$\alpha$
Füge in das Universum der Haupt- DRS einen neuen DR ein:		
		x
Suche Vorgänger-DR y, für den folgende DRS-Bedingung existiert:		
		$\alpha(y)$
IF y ex.	Einfügen in $\text{Con}_k$	x=y
ELSE	Einfügen in $\text{Con}_k$	$\alpha(x)$
Ersetze im Strukturbaum:		
	x	für
		EN
		$\alpha$

Abbildung 4.6: Konstruktionsregel KR.EN (endgültige Version).

$x_1$	$y_1$	$x_2$	$y_2$
object_4( $x_1$ )			
kriegsstraße( $y_1$ )			
$x_2 = x_1$			
ettlinger_straße( $y_2$ )			
$x_1$ enter $y_1$			
$x_2$ approach $y_2$			

Abbildung 4.7: Modifikation der DRS aus Abbildung 4.4 durch Anwenden der neuen Konstruktionsregel KR.EN (Abbildung 4.6).

Namen ‘object\_4’ trägt. Man betrachte nun, und zwar strikt in der angegebenen Reihenfolge, folgende begriffliche Beschreibung:

- (4.1) drive\_on(object\_4, the\_upper\_filling\_lane, 1).
- (4.2) stop\_on(object\_4, the\_first\_filling\_place, 2).
- (4.3) drive\_on(object\_6, the\_lower\_filling\_lane, 3).
- (4.4) stop\_on(object\_6, the\_first\_filling\_place, 4).



‘*The first filling place*’ ist keine eindeutige Beschreibung, da es an einer Tankstelle mehr als einen ersten Tankplatz gibt, denn jede Tankspur besitzt einen ebensolchen. Der ‘erste Tankplatz’ in (4.2) ist jedoch eindeutig bestimmt, da sich aus (4.1) ergibt, daß es sich um den ersten Tankplatz der oberen Tankspur handeln muß. Entgegen dem Beispiel (3.1, 3.2) ist hier die Eindeutigkeit der definiten Beschreibung nicht durch Übereinstimmung, sondern ‘lediglich’ durch einen ‘räumlichen Zusammenhang’ zwischen den definiten Beschreibungen aus (4.1, 4.2) gewährleistet. Als Fazit soll festgehalten werden, daß definite Beschreibungen stets ein bestimmtes Individuum eindeutig bezeichnen, wobei sich diese Eindeutigkeit u. U. aus einem mehr oder weniger breiten Kontext ergeben kann. Vor diesem Hintergrund kann für die begriffliche Beschreibung (4.1 - 4.4) die zugehörige DRS angegeben werden (Abbildung 4.8). Der eben detek-

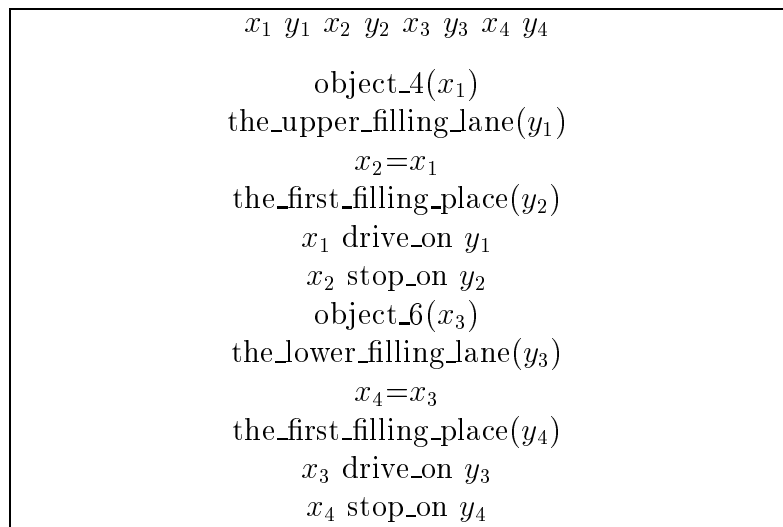


Abbildung 4.8: Zugehörige DRS zu den Fakten (4.1 - 4.4).

tierte Zusammenhang zwischen Tankspur und Tankplatz, der die Eindeutigkeit des jeweiligen ersten Tankplatzes gewährleistet, spiegelt sich in der DRS nicht unmittelbar wider, denn dazu ist Weltwissen erforderlich, das in (4.1-4.4) nicht enthalten ist. Da wir zudem, wie bereits weiter oben erwähnt, von einer zu Anfang leeren DRS ausgehen, ist diesen Weltwissen auch nicht Teil der DRS. Abbildung 4.9 gibt die Konstruktionsregel KR.DB zur Behandlung definitiver Beschreibungen an. Vergleicht man diese Konstruktionsregel mit der Konstruktionsregel zur Behandlung von Eigennamen (KR.EN, Abbildung 4.6), so fällt auf, daß sich beide Regeln ähneln, wobei KR.DB lediglich den ‘ELSE-Teil’ von KR.EN aufgreift. Diese Einschränkung hat ihre Berechtigung. Würde aufgrund des gleichlautenden Nomens in der DRS aus Abbildung 4.8 statt der Diskursbedingung ‘*the\_first\_filling\_place*( $y_4$ )’ die Diskursbedingung ‘ $y_4 = y_2$ ’ generiert, so wäre man einem inhaltlichen Interpretationsfehler aufgesessen; denn, wie bereits diskutiert, tragen die Individuen, die mit den DRen  $y_4$  bzw.  $y_2$  referenziert werden, zwar dieselben Bezeichnungen, sind jedoch nicht identisch. Es ist zu beach-

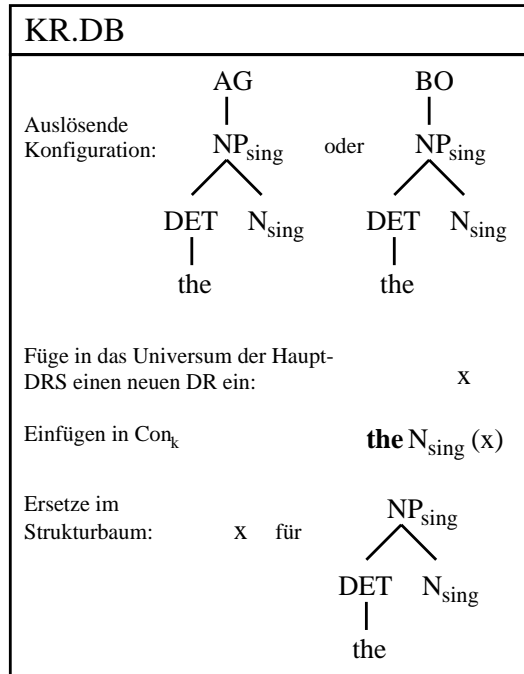


Abbildung 4.9: Konstruktionsregel KR.DB für definite Beschreibungen.

ten, daß KR.DB nur die Behandlung definiter Beschreibungen erlaubt, die im Singular stehen. Die Behandlung des Plurals folgt später. Desweiteren werden von KR.DB nur definite Beschreibungen behandelt, bei denen auf den bestimmten Artikel ‘the’ sofort ein Nomen folgt (wie z.B. bei *the\_filling\_station*). Ohne daß dies explizit in der Regel angegeben wird, sei hiermit festgestellt, daß für definite Beschreibungen, die zudem eine Ordnungszahl (wie z.B. *the\_FIRST\_filling\_place*) oder einen sonstigen Zusatz (z.B. *the\_UPPER\_filling\_lane*) enthalten, analog verfahren wird (das Ergebnis sieht man in Abbildung 4.8).

#### 4.4.4 Personalpronomen

Wie bereits in einem früheren Kapitel ausgeführt, ist in einer szenischen Beschreibung aus Protagonistensicht der Agent gleichzeitig Erzähler. Er wird seine Schilderungen demzufolge in ‘Ich-Form’ gestalten. Ein entsprechendes Faktum sieht beispielsweise wie folgt aus:

(5) enter(i, the\_filling\_station, 1).

Hierbei steht ‘i’ für das englische Personalpronomen ‘ich’. Es wird an dieser Stelle klein geschrieben, da Großbuchstaben in der angegebenen Notation für Variablennamen reserviert sind. [Kamp&Reyle 93] enthält keine Angaben darüber, wie ein Diskurs

mit Subjekt in der ersten Person zu behandeln ist. Es spricht meines Erachtens jedoch nichts dagegen, das Personalpronomen ‘**ich**’ bzw. ‘**i**’ wie einen Eigennamen zu behandeln. In beiden Fällen ist der Träger der jeweiligen Beschreibung eindeutig zu identifizieren, zumal pro Diskurs vereinbarungsgemäß nur ein fester ‘Ich-Erzähler’ auftritt. Die Konstruktionsregel zur Behandlung des Personalpronomens ‘**i**’ hat demzufolge die in Abbildung 4.10 gezeigte Form. Zu beachten ist, daß das Personalpronomen ‘**i**’

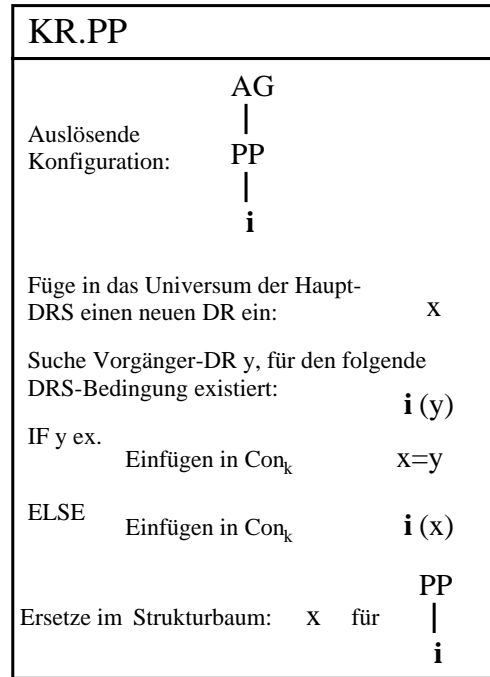


Abbildung 4.10: Konstruktionsregel KR.PP für Personalpronomen.

nur das Subjekt eines Satzes (und damit den Agenten) bezeichnen kann, nicht aber ein Bezugsobjekt. Besagtes Personalpronomen ist im übrigen das einzige, welches im Rahmen dieser Arbeit als Argument in einer begrifflichen Beschreibung auftritt. Demzufolge entfällt die Behandlung anderer Personalpronomen bei der DRS-Erstellung in der vorliegenden Arbeit.

#### 4.4.5 Prädikation

Mit *Prädikation* wird in [Kamp&Reyle 93] die Beschreibung von Eigenschaften von Individuen mit Hilfe besonderer Attribute bezeichnet. Als *Prädikat* wird in diesem Zusammenhang ein Ausdruck (in einem natürlichsprachlichen Satz) verstanden, der für eine Eigenschaft steht. Als natürlichsprachliche Prädikate kommen nach dieser Definition *Nomen, Verben, Adjektive, Präpositionen, Adverbien* und deren Kombination in Betracht. Folgende Sätze sind Beispiele für den prädikativen Gebrauch des Verbs ‘Sein’ (to be): *John is unhappy. The townhall is near the church.* [Kamp&Reyle 93] enthält

umfangreiche Ausführungen über die mögliche Behandlung von Prädikationen (Seite 260ff.). Die Semantik der in der begrifflichen Schicht extrahierten Fakten wird soweit eingeschränkt, daß wesentliche Konfliktfälle aus [Kamp&Reyle 93] hier nicht auftreten. Im Folgenden sind einige Beispiele für den prädikativen Gebrauch des Verbs ‘*to be*’ in einer begrifflichen Beschreibung angegeben:

(7.1) `be_occupied(the_first_filling_place, 1)`.

(7.2) `be_free(the_second_filling_place, 2)`.

(7.3) `be_too_close_to(object_4, another_car, 3)`.

In (7.1) liegt eine Passiv-Form des Verbs ‘*to occupy*’ vor. Dahingegen wird in (7.2) das Verb ‘*to be*’ im Zusammenhang mit einem Adjektiv verwendet. In (7.3) schließlich wird die Prädikation durch eine Adverbialphrase gebildet. Alle drei Möglichkeiten bedürfen einer eigenen Konstruktionsregel. Im Resultat werden jedoch alle drei Fälle gleich behandelt. Entsprechend der Vorgehensweise in [Kamp&Reyle 93] wird die beschriebene Eigenschaft dem gleichen DRS zugeordnet wie das Individuum, das diese Eigenschaft besitzt. Aus den Beispielen (7.1,2) resultiert folglich die in Abbildung 4.11 gezeigte DRS. Für (7.3) wird an dieser Stelle kein DRS-Eintrag angegeben, da indefinite Nomi-

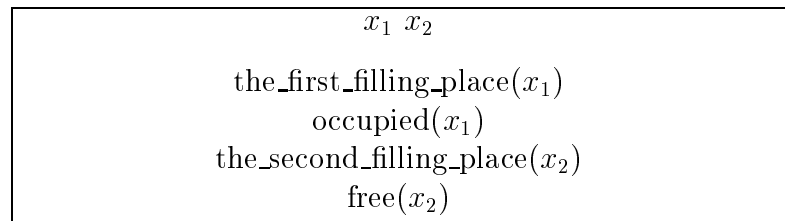


Abbildung 4.11: Zugehörige DRS zu den Fakten (7.1, 7.2).

nalphrasen wie ‘*another\_car*’ bisher noch nicht behandelt wurden. Abschließend sind zwei Konstruktionsregeln zur Behandlung von (7.1,2) angegeben (Abbildung 4.12).

#### 4.4.6 Plural

Bisher wurden lediglich Relationen behandelt, die als Argumente Bezeichnungen besitzen, die im Singular stehen. Bei der Generierung von begrifflichen Beschreibungen sind jedoch auch Objektmengen als Argumente üblich. Als erstes werden dabei solche Angaben betrachtet, welche einen natürlichsprachlichen Quantor oder eine Kardinalzahl enthalten. Man betrachte folgende Beispiele:

(8.1) `be_occupied(two_filling_places, 1)`.

(8.2) `be_occupied(all_filling_places, 2)`.

(8.3) `be_occupied(no_filling_place, 3)`.

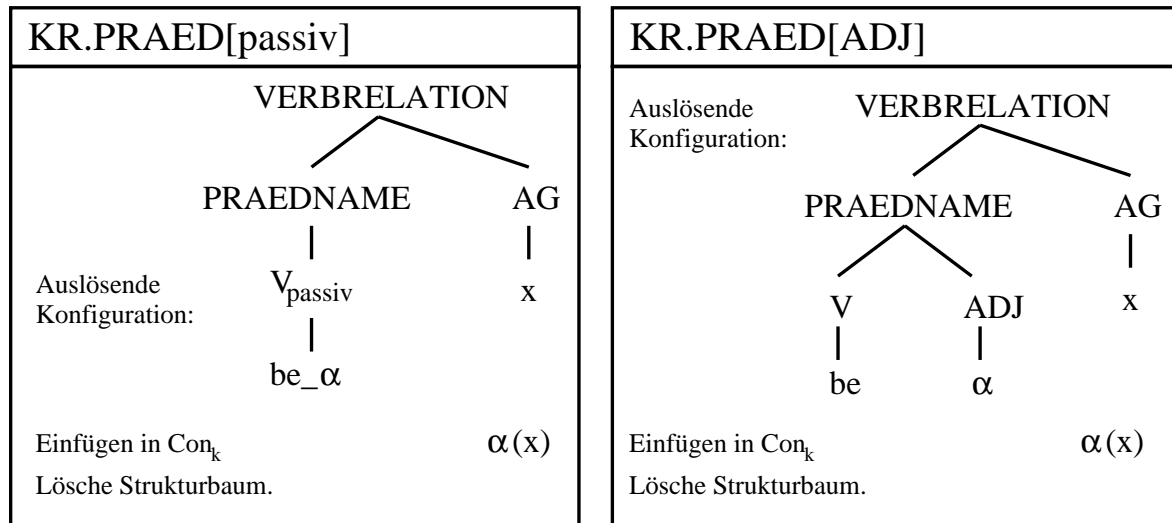


Abbildung 4.12: Konstruktionsregeln für Prädikationen.

(8.4)  $\text{quantor}(\text{many\_vehicles}, \text{on\_the\_intersection}(X), \text{come\_from}(X, \text{kriegsstra\ss e}))$ .

(8.5)  $\text{quantor}(\text{eight\_vehicles}, \text{come\_from}(X, \text{kriegsstra\ss e}), \text{drive\_to}(X, \text{ettlinger\_stra\ss e}))$ .

Es gibt zumindest zwei Möglichkeiten, die inhaltliche Aussage von Faktum (8.1) zu interpretieren und in einer DRS zu repräsentieren. Nach der ersten Lesart gibt es genau zwei Tankplätze, die beide besetzt sind. In der zweiten Lesart ist die Größe der Menge der vorhandenen Tankplätze nicht näher spezifiziert, bekannt ist allerdings, daß zwei von diesen besetzt sind. Letztgenannte ist diejenige Interpretation, die den Gegebenheiten entspricht. Durch unser zusätzliches Weltwissen um den Aufbau der Tankstelle (siehe Lageplan in Abbildung 3.2), läßt sich erstgenannte Lesart ausschließen und Faktum (8.1) im Sinne der zweiten Lesart interpretieren. Abbildung 4.13 zeigt die DRS, die nach [Kamp&Reyle 93] Faktum (8.1) in der beabsichtigten Lesart repräsentiert. Diese DRS besteht aus einer einzigen DRS-Bedingung, welche ihrerseits aus zwei (inneren) DRSen und einer Verbindungsraute besteht. DRS-Bedingungen mit diesem Aufbau werden *Duplexbedingungen* genannt. Der DR  $x$  verweist auf Vertreter des Typs *filling\_place* (linke innere DRS). Für einen bestimmten Anteil dieser Vertreter ist das Prädikat *occupied* erfüllt (rechte innere DRS). Dieser Anteil besteht aus zwei Individuen. Faktum (8.2) läßt sich analog dieser Vorgehensweise repräsentieren. Wieder handelt sich um eine in ihrer Größe nicht näher bestimmte Menge an Individuen eines

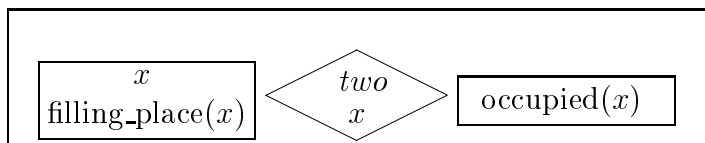


Abbildung 4.13: DRS zu Faktum (8.1).

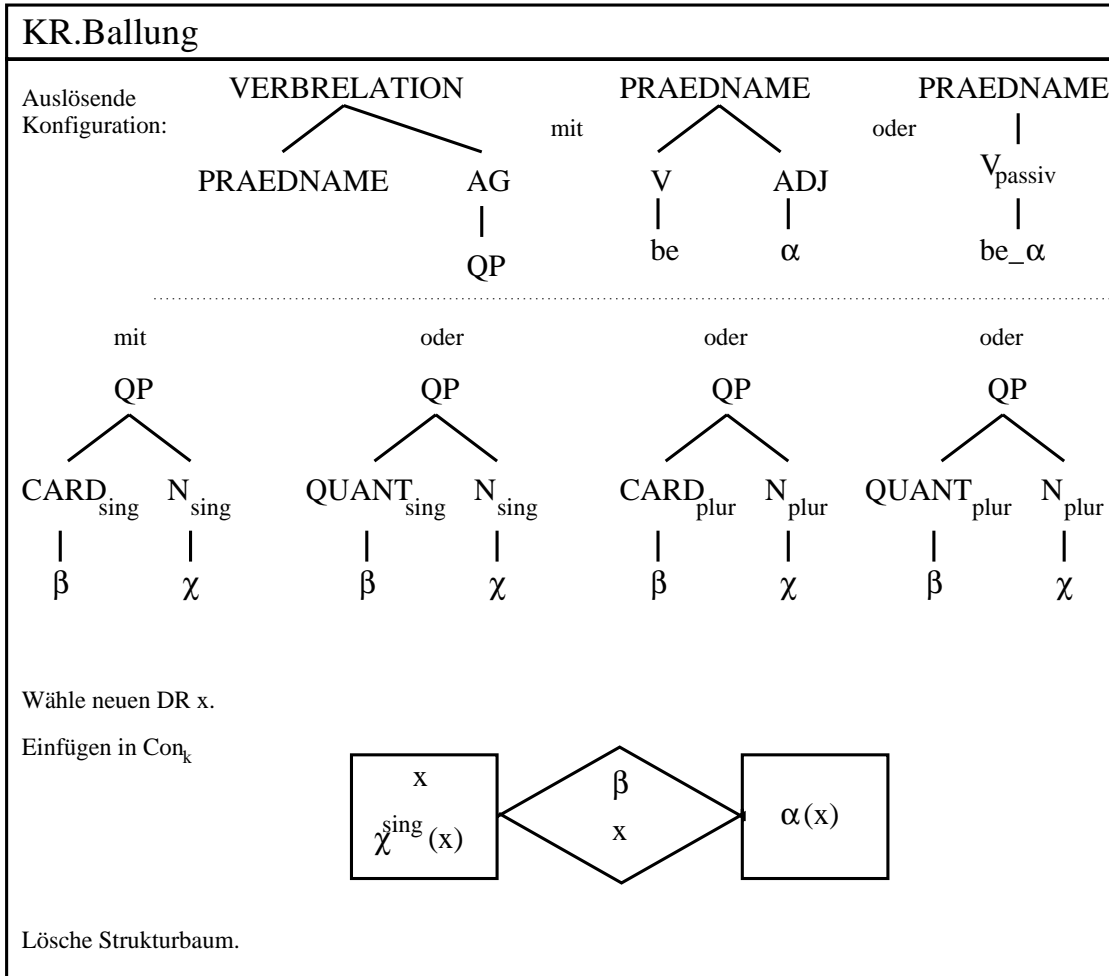


Abbildung 4.14: Konstruktionsregel KR.Ballung für Fakten der Form (8.1-3). Der Bezeichner  $\chi$  enthält Beschreibungen, die sowohl im Singular wie auch im Plural stehen können. In der Duplexbedingung wird diese Beschreibung in jedem Fall in ihre Singular-Form transformiert (angedeutet durch den hochgestellten Index *sing* bei dem betreffenden Bezeichner).

bestimmten Typs, von denen ein gewisser Anteil eine angegebene Eigenschaft erfüllt. Zur Repräsentation von (8.2) muß folglich in der DRS aus Abbildung 4.13 lediglich der Quantor *two* durch den Quantor *all* ersetzt werden. Selbst Faktum (8.3) läßt sich auf analoge Weise repräsentieren. Zur Konstruktion einer DRS des in Abbildung 4.13 gezeigten Typs wird die Konstruktionsregel aus Abbildung 4.14 verwendet. Es ist zu beobachten, daß quantorisierten Objekten entweder ein natürlichsprachlicher (relativer) Quantor QUANT voransteht (wie z.B. *few*, *some*, *many*...), oder aber eine Kardinalzahl  $CARD \in \{one, two, three, \dots\}$ . In beiden Fällen ist jedoch die Behandlung gleich. Desweiteren sind sowohl Quantoren als auch Kardinalzahlen zu unterscheiden in solche, die den Plural nach sich führen (das sind alle Zahlen größer Eins und die Mehrzahl der

relativen Quantoren), und solche, die den Singular nach sich ziehen. Zu letztgenannten gehört die Zahl Eins und der Quantor *no*, der sowohl zu den Kardinalzahlen (‘absolut Null’) als auch zu den relativen Quantoren (‘Null Prozent einer Grundmenge’) gezählt werden kann. Die Behandlung aller vier betrachteten Arten von Quantoren unterscheidet sich jedoch in der Konstruktionsregel KR.Ballung nicht. Nebenbei sei bemerkt, daß [Kamp&Reyle 93] nicht derart zwischen Kardinalzahlen und relativen Quantoren trennen. Stattdessen werden beide Wortarten, wie auch die bestimmten und unbestimmten Artikel, der Klasse der ‘Bestimmer’ (*Determiner DET*) zugeordnet (Seite 333ff.). Im Hinblick darauf, daß sich in der Bildauswertung die Berechnung relativer Quantoren und absoluter Quantoren (Kardinalzahlen) unterscheidet, wird im Rahmen der vorliegenden Arbeit diese Trennung jedoch aufrecht erhalten. [Kamp&Reyle 93] entwickeln eine sehr umfangreiche Theorie zur Behandlung von Individuenballungen, die den Plural mit sich führen. Diese Theorie behandelt hauptsächlich Mehrdeutigkeiten, die Erkennung der ‘richtigen’ Ballung und Probleme der Zuordnung des Pronomens ‘they’. Folgender Diskurs soll die Problematiken verdeutlichen (Variation verschiedener Beispiele aus [Kamp&Reyle 93]):

(9.1) *Five Lawyers hired three secretaries.*

(9.2) *Some of them are unhappy.*

Der Satz (9.1) ist vieldeutig interpretierbar. Zum einen läßt er sich so auslegen, daß insgesamt fünf Rechtsanwälte, die möglicherweise eine gemeinsame Kanzlei betreiben, zusammen drei Sekretärinnen einstellten. Eine andere Interpretation könnte besagen, daß fünf Rechtsanwälte unabhängig voneinander jeweils drei Sekretärinnen einstellten. Weiterhin könnten die fünf Rechtsanwälte unabhängig voneinander auch *in der Summe* drei Sekretärinnen eingestellt haben (z.B. einer zwei, einer eine und die restlichen drei Rechtsanwälte gar keine). Es sind noch weitere Interpretationen möglich. Jede impliziert eine andere Art der semantischen Repräsentation. Weitere Problemfelder erschließen sich, wenn man Satz (9.2) hinzufügt und versucht, die Menge zu bestimmen, die mit der Phrase ‘*some of them*’ bezeichnet werden soll. Wer ist hier unglücklich? Sind es einige der fünf Rechtsanwälte? Oder sind es einige der zwischen drei und fünfzehn Sekretärinnen? Oder sind sowohl Rechtsanwälte als auch Sekretärinnen vom Umstand des Unglücklichseins betroffen? Im Rahmen der vorliegenden Arbeit sind wir von solchen Zuordnungs- und Mehrdeutigkeitsproblemen weitestgehend befreit, da zu jedem Faktum bereits bei dessen Berechnung eine eindeutige Semantik impliziert wird, die dann durch eine entsprechend fixe Konstruktionsregel extrahiert und repräsentiert wird. Dabei wurden bisher nur Individuenballungen betrachtet, die Teil einer Prädikation sind (Faktum 8.1-3). Die Fakten (8.4) und (8.5) sind Beispiele für Ballungen, die sich aus gemeinsamen Handlungen der Individuen ergeben. Das erste Argument der Relation *quantor* gibt die Ballung an, das zweite definiert die Grundmenge GM, hinsichtlich derer der Quantor berechnet wurde, und das letzte Argument gibt die Untermenge UM an. So besagt Faktum (8.4), daß viele der Fahrzeuge, die sich auf der Kreuzung befinden (GM), aus der Kriegsstraße kamen (UM). Faktum (8.5) hingegen ist so zu in-

interpretieren, daß acht der Fahrzeuge, die aus der Kriegsstraße kamen, in die Ettlinger Straße weiterfahren.

Wir beginnen mit der Behandlung von Faktum (8.4), bei welchem die Grundmenge GM durch eine Prädikation und die Untermenge UM durch Gemeinsamkeiten in der Handlung bestimmt wird. Abbildung 4.15 enthält die DRS für die Fakten (8.4,5), die durch entsprechende Anpassung der Konstruktionsregel KR.Ballung entsteht. Diese

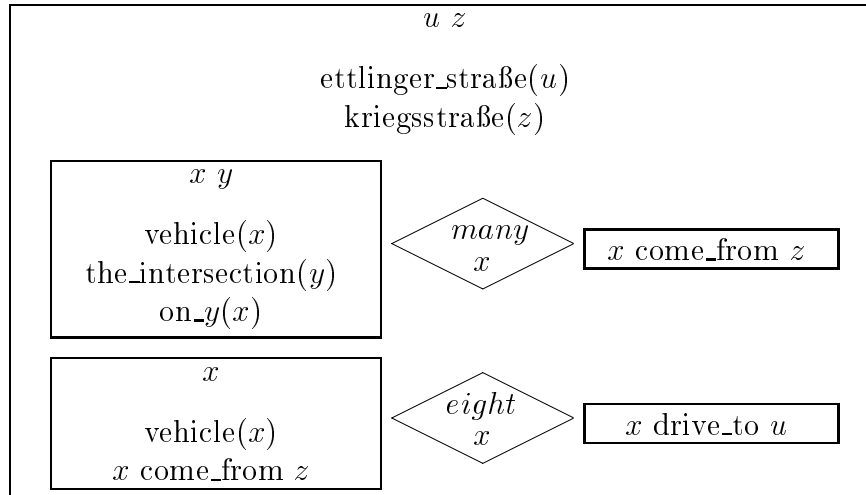


Abbildung 4.15: DRS zu den Fakten (8.4,5) (vorläufige Version).

DRS gibt die in den Fakten (8.4,5) enthaltene Semantik korrekt wieder, birgt jedoch einen Nachteil, der deutlich wird, wenn man das natürlichsprachliche Äquivalent dieser beiden Fakten betrachtet. Die beiden Fakten lassen sich wie folgt zusammenfassen: *Viele der Fahrzeuge auf der Kreuzung kommen aus der Kriegsstraße. Von diesen wiederum fahren acht in die Ettlinger Straße.* Die Übereinstimmung der Untermenge des Faktums (8.4) mit der Grundmenge des Faktums (8.5), welche in der natürlichsprachlichen Übertragung durch die Phrase *‘von diesen’* angezeigt wird, wird aus der DRS aus Abbildung 4.15 nicht deutlich. Eine alternative DRS, die nach [Kamp&Reyle 93] semantisch gleichbedeutend ist, diesen Zusammenhang aber herausstellt, ist in Abbildung 4.16 zu sehen. Im Gegensatz zu den DRSen, die mit Kleinbuchstaben bezeichnet werden, verweisen großbuchstabige DRSen auf Individuenmengen. Die Mengen  $X$  bzw.  $Y$  bezeichnen all jene Individuen, die den Bedingungen der jeweiligen inneren DRS genügen. Die Bedeutung der Duplexbedingungen der DRS aus Abbildung 4.16 erschließt sich dabei intuitiv. Es folgen in den Abbildungen 4.17 und 4.18 die Konstruktionsregeln, die aus Fakten des Typs (8.4,5) eine DRS der zuletzt angegebenen Form ableiten. [Kamp&Reyle 93] sehen Möglichkeiten vor, die Fakten (8.4,5) noch *‘kompakter’* in einer DRS zu übertragen. Man betrachte folgende natürlichsprachliche Formulierung des Faktums (8.5): *Von den Fahrzeugen, die aus der Kriegsstraße kommen, fahren acht in die Ettlinger Straße.* Dieser Satz ist gleichbedeutend mit folgendem: *Acht Fahrzeuge kommen aus der Kriegsstraße und fahren weiter in die Ettlinger Straße.*



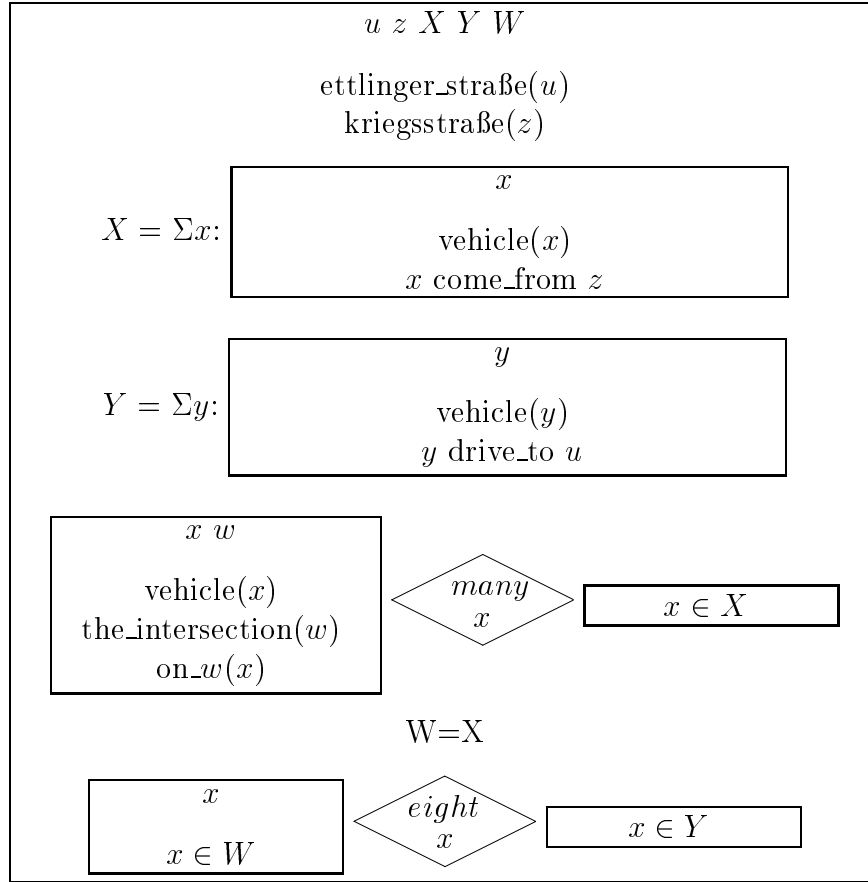


Abbildung 4.16: DRS zu den Fakten (8.4,5) (endgültige Version).

Hierbei ist zu beachten, daß die Option offen gehalten wird, daß es Fahrzeuge gibt, die zwar auch aus der Kriegsstraße kommen, aber nicht in die Ettlinger Straße fahren. Diese sind in den erwähnten acht nicht enthalten. Ausgehend von letztgenannter natürlichsprachlichen Interpretation kann Faktum (8.5) auch wie in Abbildung 4.19 als DRS repräsentiert werden. Die Diskursbedingung  $|X| = 8$  besagt, daß die Menge  $X$  an Individuen, die durch die angegebene innere DRS gebildet wird, aus acht Elementen besteht. Die Repräsentation aus Abbildung 4.19 ist verglichen mit der aus Abbildung 4.16 wesentlich kompakter. Es wird zumindest eine Duplexbedingung eingespart. Andererseits birgt die alternative Darstellung auch einige Nachteile in sich, so daß die Darstellung aus Abbildung 4.16 trotz des größeren formalen ‘Aufwands’ beibehalten wird. Zum einen verhindert die kompakte Darstellung, daß Grundmengen und Untermengen verschiedener quantorisierten Fakten zueinander in Beziehung gesetzt werden können (was sich ja in Abbildung 4.16 als zweckmäßig erwiesen hatte). Zweitens kann die kompakte Darstellung nur für Kardinalzahlen, nicht aber für die meisten relativen Quantoren angewandt werden, wodurch die Gleichbehandlung von relativen Quantoren und Kardinalzahlen aufgehoben werden würde. Drittens stimmen

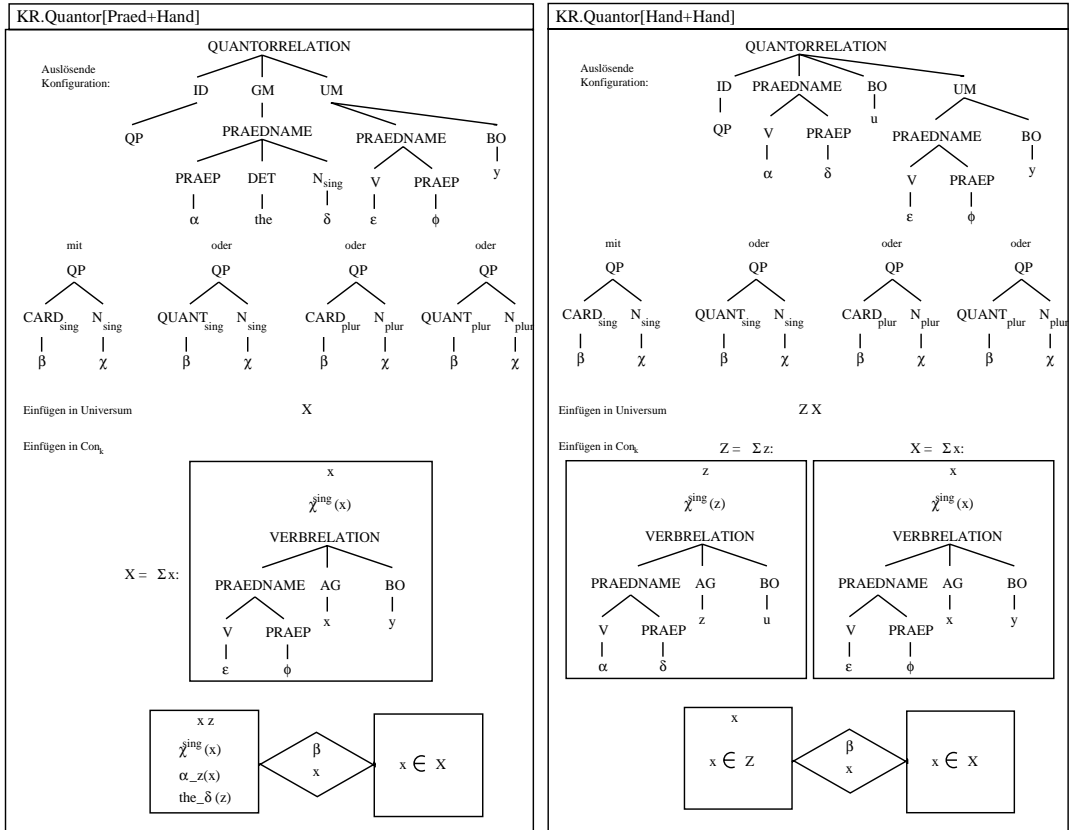


Abbildung 4.17: Konstruktionsregeln für Fakten der Form (8.4).

Duplexbedingungen in ihrem Aufbau recht anschaulich mit der verwendeten Notation zur Beschreibung von Fakten überein. Die Fakten (8.4, 8.5) enthalten drei Argumente, nämlich die Individuenballung mit Quantor, die Grundmenge und die Untermenge. Dementsprechend enthält auch jede Duplexbedingung drei Bestandteile: eine innere DRS, die die Grundmenge angibt; eine weitere DRS, die die Untermenge repräsentiert und die ‘Verbindungsraute’, die den Quantor angibt. Aufgrund dieser formalen Übereinstimmung erscheint es angebracht, Fakten der Form (8.4, 8.5) auch weiterhin in der etwas ‘aufwendigeren’ Form aus Abbildung 4.16 zu repräsentieren.

### 4.4.7 Indefinite Beschreibungen

*Definite* Beschreibungen von Individuen wurden bereits in einem der vorangegangenen Abschnitten behandelt (zumindest solche, die im Singular stehen). Darunter wurden Individuenbeschreibungen verstanden, die mit dem bestimmten Artikel eingeleitet werden. Durch diese Art der Beschreibung wird ein bestimmtes Individuum eindeutig festgelegt. *Indefinite* Beschreibungen, um die es in diesem Abschnitt gehen soll, legen jedoch Individuen nicht eindeutig fest, sondern geben lediglich deren Typ an.

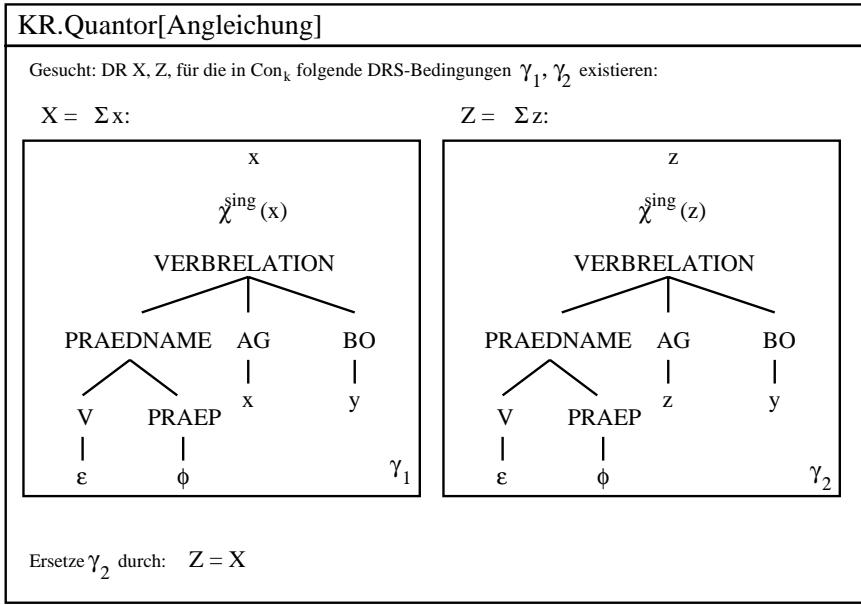


Abbildung 4.18: Konstruktionsregeln für Fakten der Form (8.4), Fortsetzung.

Sie werden meist durch einen indefiniten Bezeichner, welches zum Beispiel ein unbestimmter Artikel sein kann, eingeleitet. Im folgenden sind einige Beispiele für indefinite Beschreibungen angegeben:

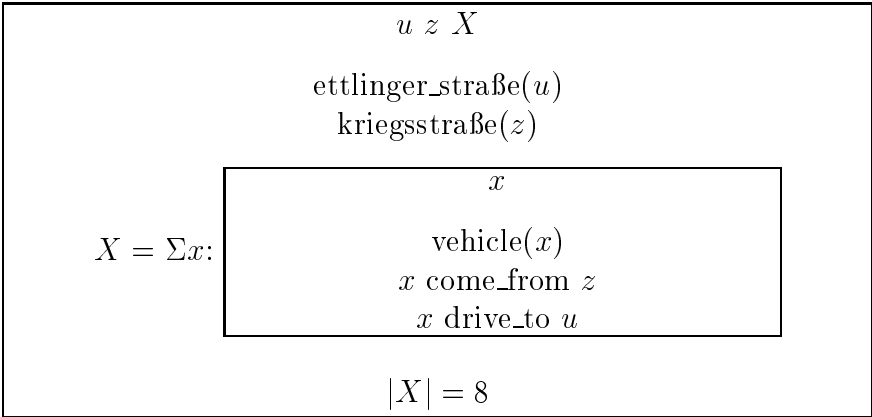


Abbildung 4.19: Alternative DRS zum Faktum (8.5).

- (10.1) `overtake(object_4, a_car, 1)`.
- (10.2) `overtake(object_4, another_car, 2)`.
- (10.3) `overtake(object_4, other_cars, 3)`.
- (10.4) `leave_enough_space_for(object_4, other_cars, 4)`.
- (10.5) `take_in(object_4, petrol, 5)`.

Faktum (10.1) stellt ein einfaches Beispiel für eine indefinite Beschreibung dar. Das definite Individuum ‘object\_4’ überholt ein Individuum, über dessen genaue Identität keine Hinweise gegeben sind, außer dem, daß es einem bestimmten Typ, nämlich dem Typ ‘Fahrzeuge’, angehört. Im vorigen Abschnitt wurden schon, damals allerdings ohne näher darauf einzugehen, solche Typzuordnungen verwendet. Man betrachte beispielsweise noch einmal Abbildung 4.16. In dieser DRS tritt mehrfach die DRS-Bedingung  $vehicle(x)$  auf. Diese Bedingung verlangt, daß das Individuum, das durch den entsprechenden DRen repräsentiert wird, vom Typ ‘vehicle’ ist. Sie stellt folglich eine indefinite Beschreibung dar, die dadurch zustande kam, daß von der Quantorisierungsphrase QP (siehe Konstruktionsregel in Abbildung 4.17) der jeweilige Quantor abgetrennt wurde. Analog dazu wird Faktum (10.1) in die DRS aus Abbildung 4.20 übertragen. Diese

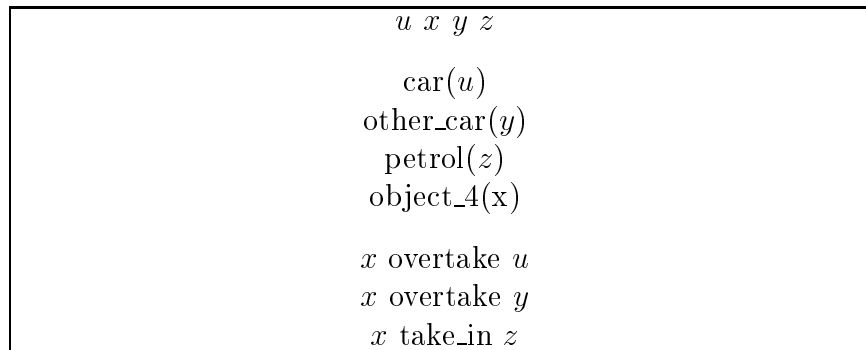


Abbildung 4.20: DRS zu den Fakten (10.1,2,5).

DRS enthält zudem die Repräsentation der Fakten (10.2) und (10.5). Betrachtet man Faktum (10.2) für sich allein, so ist die Phrase ‘another\_car’ als ‘anderes Fahrzeug als Object\_4 selbst’ zu deuten. Es ist zwar klar, daß ein Fahrzeug nicht sich selbst überholen kann, trotzdem könnte man geneigt sein, das Vorhandensein des Zusatzes ‘other’ dazu zu nutzen, diese Abgrenzung in der DRS explizit zu dokumentieren, etwa durch die Zusatzbedingung  $x \neq y$ . Um eine solche Bedingung jedoch extrahieren zu können, muß das Wissen explizit vorhanden sein, daß es sich bei ‘Object\_4’ um ein Fahrzeug handelt. Dies ist in der angegebenen DRS nicht vorhanden. Deshalb wird die Phrase ‘another\_car’ genauso behandelt, wie in Abbildung 4.20 dargestellt. Es ist dabei zu beachten, daß nach [Kamp&Reyle 93] der unbestimmte Artikel nicht in die DRS mitgeht. Bei dem Argument ‘petrol’ in Faktum (10.5) handelt es sich ebenfalls um eine indefinite Beschreibung. Es ist kein bestimmtes Benzin gemeint, sondern lediglich

der Typ. Flüssigkeiten wie Benzin oder Wasser gehören zu den nicht zählbaren Dingen, so daß sie ohne den unbestimmten Artikel ‘a’ stehen können. Ansonsten erfolgt die Behandlung von (10.5) analog zu (10.1). Die Fakten (10.3,4) enthalten indefinite Beschreibungen im Plural. Eine Eigenart solcher Beschreibungen ist, daß sie zwar im Plural stehen, aber nicht notwendigerweise, wie wohl zu erwarten wäre, auf eine Individuenmenge mit mehr als einem Element verweisen müssen. Aus (10.3) läßt sich ableiten, daß mehr als ein Fahrzeug überholt wurde. Dahingegen sagt (10.4) zwar aus, daß Platz für andere Fahrzeuge frei gelassen wurde, es sagt aber nichts darüber aus, wie viele andere Fahrzeuge diesen Platz genutzt haben. Mit anderen Worten: Die Menge der ‘anderen Fahrzeuge’, die Nutzen aus dem freien Platz ziehen konnten, kann im Extremfall sogar leer sein, und dennoch ist die Verwendung des Plurals korrekt. Selbst wenn nur Platz für *genau ein* Fahrzeug gelassen wurde, ist die Verwendung des Plurals korrekt, da unbestimmt bleibt, welches Fahrzeug, wenn überhaupt, diesen Platz nutzen wird. Es könnte insbesondere der Fall eintreten, daß wiederum mehrere Fahrzeuge den Platz nutzen, und zwar nicht gleichzeitig, sondern nacheinander. Abbildung 4.21 enthält die Konstruktionsregel KR.IB zur Behandlung indefiniter Beschreibungen mit unbestimmtem Artikel im Singular und ohne Artikel und Quantoren im Plural.

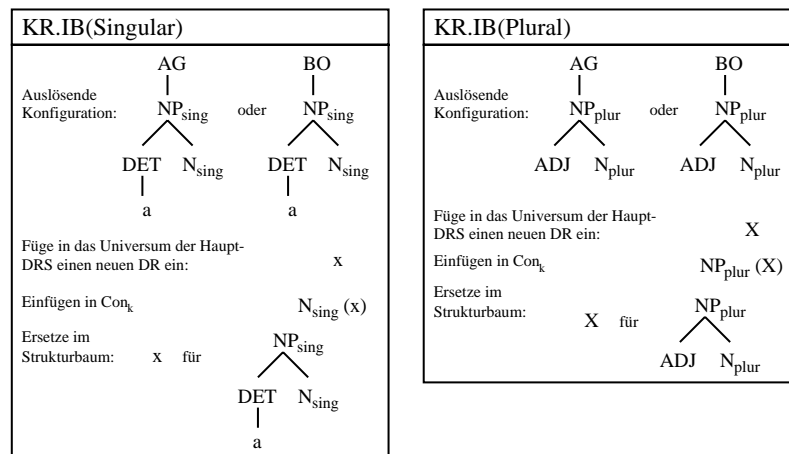


Abbildung 4.21: Konstruktionsregeln für indefinite Beschreibungen. Der wesentliche Unterschied bei der Behandlung von Singular und Plural liegt in der Wahl des DRen. Im Singular wird ein kleinbuchstabiger DR verwendet, während im anderen Fall ein großbuchstabiger zur Repräsentation von Individuenmengen verwendet wird.

#### 4.4.8 Zeit

Die bisher entwickelten Konstruktionsregeln berücksichtigen nicht die temporalen Informationen, die sich aus einer begrifflichen Beschreibung ableiten lassen. Anhand der nun folgenden Beispiele wird dargelegt, wie *Abläufe* in DRSen repräsentiert werden. Dabei werden wieder englischsprachige Bezeichnungen verwendet, da auf eine englisch-

sprachige Beschreibung abgezielt wird (zur Begründung sei auf die einleitenden Worte zum zweiten Kapitel der vorliegenden Arbeit zurück verwiesen).

(12.1) 100 : 200 ! enter(object\_5, the\_filling\_station, 1).

(12.2) 100 : 300 ! stand\_on(object\_6, the\_first\_filling\_place, 2).

(12.3) 201 : 300 ! overtake(object\_5, object\_6, 3).

Bei der Repräsentation von Abläufen wird im folgenden ein DR  $a$  verwendet. Der Ablaufzeitraum  $a$  wird durch zusätzliche Angaben, die sich aus dem Faktum selbst oder aus dessen Kontext ergeben, eingeschränkt. So könnte eine DRS, die zusätzlich die temporalen Aspekte von (12.1-3) berücksichtigt, das in Abbildung 4.22 gezeigte Aussehen haben. Der jeweilige Ablaufzeitraum  $a_i$  wird durch die Grenzen  $t_{i1}$  bzw.  $t_{i2}$

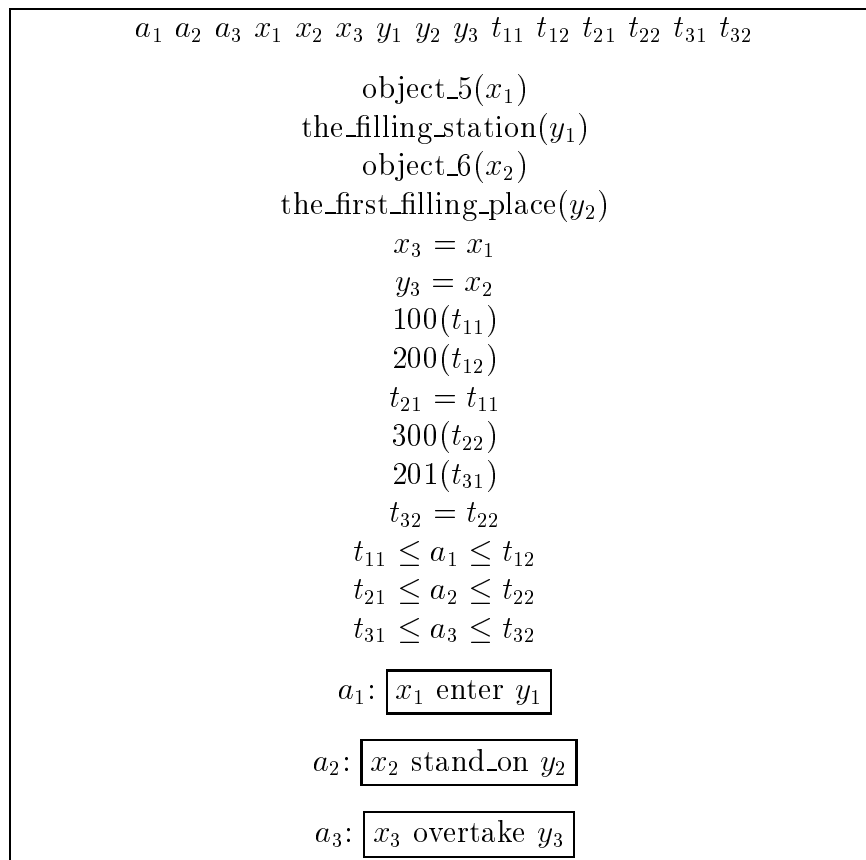


Abbildung 4.22: DRS zu den Fakten (12.1-3) inklusive temporalen Aspekte (vorläufige Version).

eingegrenzt, deren Definition sich aus dem Gültigkeitsintervall des jeweiligen Faktums ergibt. Das Zeitmodell, das aufgrund dieser Vorgehensweise in der DRS aus Abbildung 4.22 zum Tragen kommt, entspricht demjenigen aus [Allen 83]. Dieses intervallbasierte

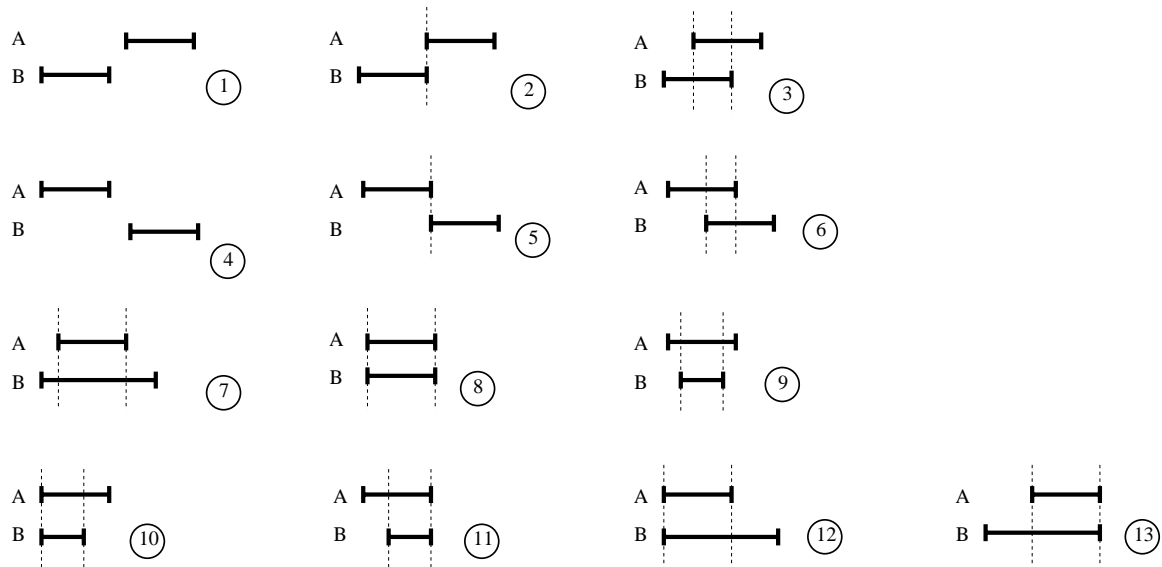


Abbildung 4.23: Dreizehn Möglichkeiten der zeitlichen Konstellation von Ereignissen nach [Allen 83].

Zeitmodell sieht dreizehn Konstellationen vor, wie sich zwei Ereignisse zeitlich zueinander verhalten können (Abbildung 4.23). Wie bereits in der Einleitung zu diesem Kapitel dargelegt, findet das Zeitmodell von [Allen 83] bei [Kamp&Reyle 93] keine Anwendung. Und auch für die vorliegende Arbeit erweist es sich nicht als ideal. Man betrachte dazu folgende begrifflichen Beschreibungen:

- (12.4) 100 : 200 ! enter(object\_5, the\_filling\_station, 4).  
 201 : 300 ! arrive\_at(object\_6, the\_first\_filling\_place, 5).
- (12.5) 100 : 200 ! enter(object\_5, the\_filling\_station, 4).  
 200 : 300 ! arrive\_at(object\_6, the\_first\_filling\_place, 5).
- (12.6) 100 : 200 ! enter(object\_5, the\_filling\_station, 4).  
 199 : 300 ! arrive\_at(object\_6, the\_first\_filling\_place, 5).

Die drei begrifflichen Beschreibungen (12.4-6) unterscheiden sich nur durch die Halbbildnummer, mit der der jeweils zweite Ablauf beginnt. Der minimale Unterschied von einem Halbbild, also einer fünfzigstel Sekunde, führt dazu, daß nach [Allen 83] alle drei begrifflichen Beschreibungen einer anderen Ereigniskonstellations angehören. So stellt (12.4) ein Vertreter der Konstellation (4) in Abbildung 4.23 dar, (12.5) korrespondiert zu (5) und (12.6) zu (6). Im Ergebnis bedeutet dies, daß eine minimale zeitliche Änderung, die jenseits der normalen menschlichen Wahrnehmung liegt und dazu noch einem Meßrauschen unterliegen kann, zu unangemessener Signifikanz bei der semantischen Repräsentation führen würde. Das ereignisbasierte Zeitmodell von

[Kamp&Reyle 93] hingegen kommt, wie bereits an anderer Stelle ausgeführt, mit den zwei Relationen ‘ $\circ$ ’ für zeitliches Überlappen und ‘ $<$ ’ für zeitliches Aufeinanderfolgen aus. Dabei werden die Fälle (3,6-13) aus Abbildung 4.23 dem zeitlichen Überlappen zugeordnet, die Fälle (1,2,4,5) hingegen dem zeitlichen Aufeinanderfolgen. Beim Übergang vom intervallbasierten Zeitmodell nach [Allen 83] zum ereignisbasierten Zeitmodell nach [Kamp&Reyle 93] wird die DRS aus Abbildung 4.22 in die DRS aus Abbildung 4.24 überführt.

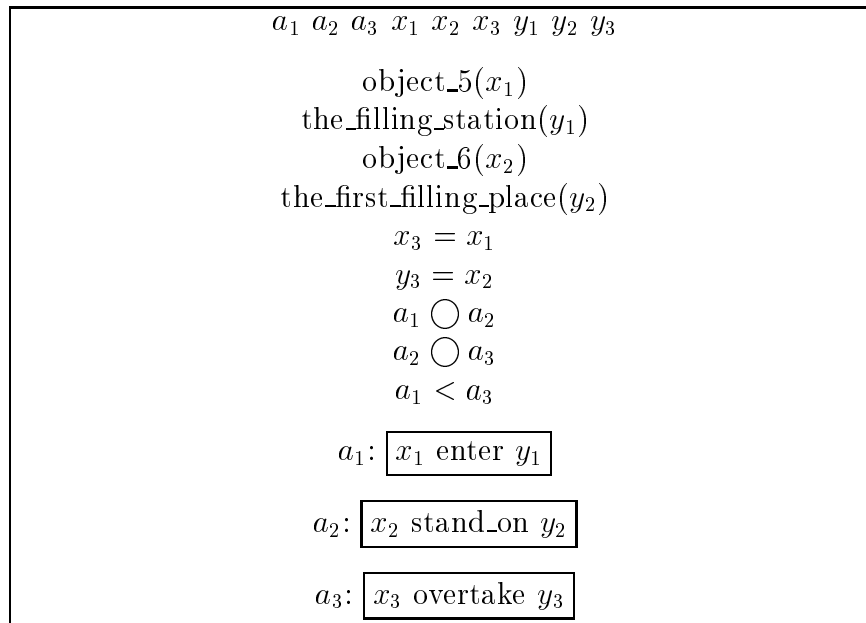


Abbildung 4.24: DRS zu den Fakten (12.1-3) inklusive temporaler Aspekte, endgültige Version.

#### 4.4.9 Weitere Bestandteile von Beschreibungen

In den vorausgegangenen Abschnitten wurde dargestellt, wie bestimmte Bestandteile begrifflicher Beschreibungen (wie z.B. indefinite oder definite Beschreibungen sowie quantorisierte Beschreibungen) durch Verwendung der DRT nach [Kamp&Reyle 93] in die semantische Repräsentationsform DRS übertragen werden. Die bisher generierten begrifflichen Beschreibungen überschreiten jedoch bereits jetzt in ihrer Ausdruckskraft das von besagten Autoren behandelte Spektrum. Die entsprechenden ‘extraordinären’ Fakten umfassen meist Relationen, die Beziehungen zwischen Ereignissen beschreiben (*Metarelationen*). Andere Beispiele sind sprachliche Einschränkungen, die Schlußfolgerungen aus Beobachtungen als Vermutungen des Beobachters kennzeichnen. Dies kann durch entsprechende Beschreibungszusätze wie z.B. durch das Adverb *presumably* geschehen, welche den vermutenden Charakter einer Schlußfolgerung ausdrückt.



Darüber hinaus sind noch weitere Beschreibungszusätze möglich, die in der begrifflichen Beschreibung bereits extrahiert werden, jedoch derzeit nicht in einer DRS repräsentiert werden können, da sie in [Kamp&Reyle 93] nicht behandelt wurden. Somit finden diese Beschreibungszusätze auch im Rahmen der vorliegenden Arbeit keinen Einzug in die DRS, werden jedoch bei der Erzeugung natürlichsprachlicher Beschreibung weiter berücksichtigt.

## 4.5 Ableitung von Texten aus der DRS

In den vorausgegangenen Abschnitten wurde beschrieben, wie aus einer automatisch erstellten begrifflichen Beschreibung eine rechnerinterne semantische Repräsentation in Form einer DRS erstellt wird. Im aktuellen Abschnitt soll nun erörtert werden, wie aus einer DRS ein natürlichsprachlicher Text erstellt werden kann. Ein solches Vorhaben wirft eine Fülle von Grundsatzfragen auf, deren erschöpfende Behandlung weit über den Rahmen einer einzigen Dissertation hinausgeht. [Kamp&Reyle 93] gehen in ihren Ausführungen nicht darauf ein, wie aus einer DRS ein Text abgeleitet werden kann, sondern beschränken sich allein auf die umgekehrte Zielsetzung. Demzufolge kann im Rahmen der vorliegenden Arbeit keine erschöpfende Methodik zur Transformation von DRSen in Text vorgelegt werden, einige Überlegungen in dieser Hinsicht erscheinen jedoch möglich.

### 4.5.1 Grundsätzliche Überlegungen

Das zu behandelnde Problem der Transformation von DRSen in natürlichsprachlichen Text wird im folgenden schematisch beleuchtet. Es stellt sich zunächst einmal wie folgt dar:

**Gegeben:** DRS.  
**Gesucht:** Methodik, wie aus dieser DRS ein natürlichsprachlicher Text abgeleitet werden kann.  
**Ergebnis:** Text.

Diese Problemformulierung bedarf der Präzisierung. Aus einer DRS soll genau ein Text entstehen, der die Information enthält, die auch in der DRS enthalten ist. Andererseits wurde schon zu Beginn dieser Arbeit festgestellt, daß ein und derselbe Ablauf auf sehr unterschiedliche Weise beschrieben werden kann. Zu diesem Zweck wurde der Begriff ‘Perspektive einer Beschreibung’ eingeführt. Abhängig von der jeweiligen Perspektive wird eine begriffliche Beschreibung und daraus eine DRS berechnet. Das Bestreben, eine Beschreibung zu extrahieren, die eine bestimmte, benutzerdefinierte Perspektive einnimmt, wird im folgenden als ‘*Erreichen eines Erzählziels*’ bezeichnet. Demzufolge soll auch der aus einer DRS abzuleitende Text das angestrebte Erzählziel erreichen. Das

Problem der Transformation von DRSen in natürlichsprachlichen Text wird demzufolge wie folgt modifiziert:

<b>Gegeben:</b>	Erzählziel.
<b>Gegeben:</b>	Begriffliche Beschreibung, die hinsichtlich dieses Erzählziels erstellt wurde.
<b>Gegeben:</b>	DRS, die aus der begrifflichen Beschreibung abgeleitet wird.
<b>Gesucht:</b>	Methodik, wie aus dieser DRS ein natürlichsprachlicher Text abgeleitet werden kann.
<b>Ergebnis:</b>	Text, der das gegebene Erzählziel erreicht.

Ein großes Problem bei der Texterstellung aus einer DRS stellt die Frage, wie eine sinnvolle Textstruktur erzielt werden kann. Soll der Text wirklich jede Information aus der DRS umsetzen? Wie kann vermieden werden, daß etwaig vorhandene redundante Information aus der DRS sich als Redundanz im Text auswirkt? Wie erzielt man sukzessive eine 'sinnvolle' Textstruktur? Was ist eine sinnvolle Textstruktur? Dieses Kaleidoskop an Fragestellungen ist längst nicht vollständig, wenn es um die Texterzeugung aus einer semantischen Repräsentation geht. Und es wird im Rahmen dieser Arbeit auch nicht gelöst werden können. Allerdings lassen sich die angegebenen Problemstellungen einschränken. Es wird davon ausgegangen, daß die gegebene DRS dem gegebenen Erzählstil genügt, sie somit alle relevanten Informationen enthält, die zum Erreichen des Erzählziels benötigt wird. Zum anderen wird davon ausgegangen, daß sie nur diese Informationen enthält und keine weiteren (also frei von Redundanz ist). Redundanzfreie, das Erzählziel erfüllende DRSen werden im folgenden *pragmatische DRSen* genannt. Für pragmatische DRSen sind die angesprochenen Probleme der Fokussierung und der Redundanzvermeidung nicht mehr relevant. Für pragmatische DRSen genügt es demnach, eine DRS-interne Struktur, d.h. eine bestimmte Abarbeitungsreihenfolge festzulegen, hinsichtlich derer die Textstruktur erstellt wird. Somit ist die Transformation pragmatischer DRSen in Text eindeutig und vollständig. Der entsprechende Auswertungsschritt wird im folgenden Abschnitt vorgestellt. Es ergibt sich folgende Problemspezifikation:

<b>Gegeben:</b>	Erzählziel.
<b>Gegeben:</b>	Begriffliche Beschreibung, die alle relevante (und nur diese) Information zur Erfüllung dieses Erzählziels enthält.
<b>Gegeben:</b>	Pragmatische DRS mit festgelegter Abarbeitungsstruktur.
<b>Gesucht:</b>	Menge von Ableitungsregeln zur Ableitung von Text aus der DRS.
<b>Ergebnis:</b>	Text, der das gegebene Erzählziel erreicht.

Die Bestimmung eines Erzählziels und die Berechnung begrifflicher Beschreibungen abhängig von gegebenen Erzählzielen wurde bereits in früheren Abschnitten diskutiert. Die Ableitung der DRS aus einer begrifflichen Beschreibung war Thema des

vorangegangenen Abschnitts. Die Frage, wie die Abarbeitungsstruktur einer pragmatischen DRS festgelegt wird, wird im folgenden Abschnitt geklärt. Eine pragmatische DRS läßt sich grundsätzlich auf zwei Arten erzeugen: Zum einen, indem bereits bei der Erzeugung der DRS aus der begrifflichen Beschreibung auf die Einhaltung dieses Kriteriums geachtet wird. Dies ist die Vorgehensweise, die bisher im Rahmen dieser Arbeit gewählt wird. Zum anderen kann man Methoden vorsehen, wie aus einer nicht-pragmatischen DRS eine pragmatische DRS abgeleitet wird. Auf diese Möglichkeit wird im Fortlauf dieser Arbeit unter dem Stichwort der logischen Manipulation von DRSen und im Rahmen des Ausblicks noch eingegangen.

#### 4.5.2 Text- und DRS-Aufbau

Gegeben ist ein Erzählziel. Gegeben ist ferner eine pragmatische DRS, also eine DRS, die alle relevanten Informationen (und nur diese) zur Gewährleistung des Erzählziels enthält. Zur Transformation dieser DRS in einen natürlichsprachlichen Text muß eine DRS-interne Struktur festgelegt werden, aus der sich eine Abarbeitungsreihenfolge für die Texterzeugung ableiten läßt. Aus dieser DRS-internen Struktur kann dann mit Hilfe eines Regelwerks (das noch zu konstruieren ist) eine Textstruktur zur Generierung der natürlichsprachlichen Beschreibung entwickelt werden.

Ein natürlichsprachlicher Satz besteht in seiner grundlegenden Form aus einem Subjekt, einem Prädikat und ggf. aus einem oder mehreren Bezugsobjekten. Sätze beschreiben gewöhnlich Abläufe. Auch DRSen enthalten Ablaufbeschreibungen. Je nachdem, ob es sich um ein intransitives oder ein transitives Verb handelt, haben DRS-Bedingungen, die einen Ablauf beschreiben, folgenden Aufbau:

$$(14.1) \quad a_i: \boxed{x_i \text{ Verbalphrase}} \quad \text{oder} \quad a_i: \boxed{x_i \text{ Verbalphrase } y_i}$$

Dabei repräsentiert der DR  $a_i$  den Ablauf und damit auch die Ablaufzeit. Zwischen verschiedenen Ablaufbeschreibungen in einer DRS wurden, wie bereits bekannt, die temporalen Relationsbezeichnungen ‘ $\circ$ ’ für zeitliches Überlappen und ‘ $<$ ’ für zeitliches Aufeinanderfolgen eingeführt. Ablaufbeschreibungen lassen sich folglich chronologisch ordnen. Hierzu werden Zeitgraphen verwendet.

Ein *Zeitgraph* ist ein Graph, dessen Knoten Ablaufbeschreibungen, also DRS-Bedingungen der in (14.1) angegebenen Form, und dessen Kanten zeitliche Konstellationen zwischen Abläufen repräsentieren. Folglich gibt es zwei Arten von Kanten: *Parallelkanten* verweisen auf zeitlich überlappende Abläufe, wohingegen *Nachfolgekanten* auf zeitlich nachfolgende Abläufe verweisen. Der rekursive Aufbau dieses Graphen beginnt mit einem frühesten Ablauf, also einem Ablauf  $a_i$ , für den kein weiterer Ablauf  $a_j$  mit zugehöriger DRS-Bedingung  $a_j < a_i$  existiert. Für jeden neu einzuordnenden Ablauf  $a_k$  werden abhängig vom Vorhandensein von DRS-Bedingungen der Form ‘ $a_j < a_k$ ’ bzw. ‘ $a_j \circ a_k$ ’ Nachfolge- bzw. Parallelkanten zwischen  $a_k$  und allen in Frage kommen-

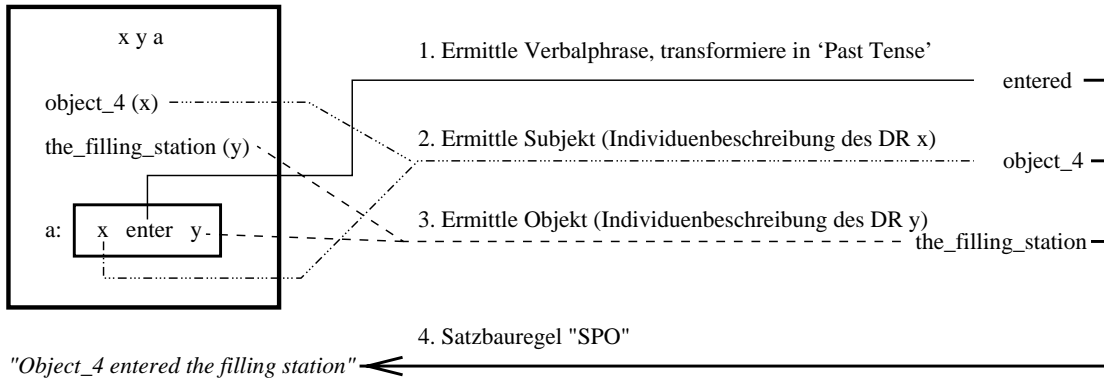


Abbildung 4.25: Satzbauregel zur Bildung eines einfachen Hauptsatzes.

den Knoten  $a_j$  des Graphen extrahiert. Durch den Aufbau von Zeitgraphen erhalten Ablaufbeschreibungen einer DRS eine Abarbeitungsreihenfolge. Diese chronologische Struktur wird als Textstruktur übernommen. Abläufe werden folglich im zu generierenden natürlichsprachlichen Text chronologisch geschildert. Damit wird die Forderung erfüllt, daß in den Erzählstilen ‘Szenische Beschreibung aus Beobachtersicht’ und ‘Szenische Beschreibung aus Protagonistensicht’ Abläufe zwingend chronologisch geschildert werden. Für den Erzählstil ‘Bericht’ ist eine chronologische Ablaufschilderung ebenfalls erlaubt. Somit ist festgelegt, in welcher Reihenfolge Ablaufbeschreibungen in natürlichsprachliche Sätze umgesetzt werden. Jede Ablaufbeschreibung entspricht zunächst einmal genau einem Satz. Es ist nun noch zu klären, wie eine aktuelle Ablaufbeschreibung im Zeitgraph in einen Satz übertragen wird. In den folgenden Unterabschnitten werden die wichtigsten in dieser Arbeit verwendeten Regeln hergeleitet. Über weitere Regeln gibt der Anhang Aufschluß.

### 4.5.3 Einfache Hauptsätze

Gegeben sei die in Abbildung 4.25 gezeigte DRS. Sei  $a$  der aktuell in einem natürlichsprachlichen Satz zu formulierende Ablauf des Zeitgraphen. Zur Satzbildung wird von der einfachsten englischen Satzbauregel ausgegangen, die besagt, daß im Satz das Subjekt an erster Stelle steht, vor dem Prädikat und dem Objekt (*SPO-Regel*). Dieser Regel wird in der Ablaufbeschreibung aus Abbildung 4.25 bereits genüge getan. Die Ablaufbeschreibung enthält zum einen ein Verb (*‘enter’*). Zum anderen verweist der erste DR  $x$  auf das Agens des Ablaufs. Dieser Agens wird als Subjekt des zu bildenden Satzes aufgefaßt. Für den DR  $x$  muß eine DRS-Bedingung existieren, die diesem DR eine Bezeichnung zuführt. Im angegebenen Fall ist es ein Eigenname (*‘object\_4’*). Das Objekt des zu bildenden Satzes läßt sich ebenfalls leicht bestimmen. Es ergibt sich aus dem DR  $y$ , welcher durch eine definite Beschreibung (*‘the\_filling\_station’*) bezeichnet wird. Somit entsteht aus der DRS in Abbildung 4.25 folgender Satz: *Object\_4 entered the filling station*. Die Vergangenheitsform wird generell verwendet. Sie ist für den Erzähl-

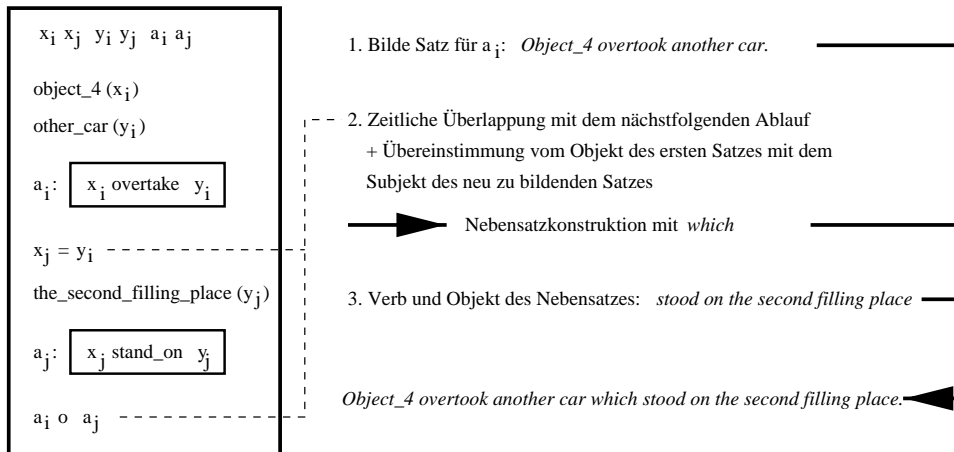


Abbildung 4.26: Satzbauregel zur Bildung eines Relativsatzes.

stil ‘Bericht’ obligatorisch, da ein Bericht einen bereits vergangenen Ablauf beschreibt. Szenische Beschreibungen aus Beobachter- und Protagonistensicht können zwar auch im Präsens formuliert werden, besonders wenn Abläufe zeitgleich zu ihrem Geschehen beschrieben werden. [Kamp&Reyle 93] schließen diese Verwendung des Präsens als ‘Reportive Speech’ aus verschiedenen Gründen in ihren Betrachtungen aus, so daß auch in diesen Fällen die Vergangenheitsform verwendet wird. Dies entspricht dem Szenario, daß ein Beobachter bzw. ein Protagonist einen vergangenen Ablauf aus seiner Sicht rekapituliert.

#### 4.5.4 Relativsätze

Abbildung 4.26 zeigt eine Möglichkeit, wie sich eine bestimmte Konstellation an DRS-Bedingungen in eine Hauptsatz-/Relativsatzkonstruktion übertragen läßt. Gegeben seien zwei sich zeitlich überlappende Abläufe, wobei das Bezugsobjekt des einen Ablaufs mit dem Agens des anderen identisch ist (DRS-Bedingung  $x_j = y_i$ ). Die Individuenbezeichnung, die in beiden Abläufen vertreten ist, wird im folgenden *Kernindividuenbezeichnung* genannt. Im Beispiel handelt es sich dabei um eine indefinite Beschreibung des Individuums (‘other car’). Statt beide Abläufe nach der ‘SPO’-Regel in zwei getrennten Sätzen zu formulieren, wird zur Erweiterung der stilistischen Vielfalt der natürlichsprachlichen Beschreibung in der angegebenen Konstellation eine Relativsatzkonstruktion angestrebt. Der Ablauf, bei welchem das Bezugsobjekt Kernindividuum ist, wird entsprechend der SPO-Regel gebildet. Der andere Ablauf wird durch einen Relativsatz beschrieben. Dabei wird das entsprechende Relativpronomen an das Bezugsobjekt des Hauptsatzes angehängt. Prädikat und Bezugsobjekt werden wie gewohnt gebildet. Es sind zur Hauptsatz-/Nebensatzkonstellation zwei zusätzliche lexikalische Informationen notwendig. Zum einen muß zur Bestimmung des passenden Relativpronomens (hier: ‘which’, statt etwa ‘who’) das Geschlecht des Kernindividuum bekannt

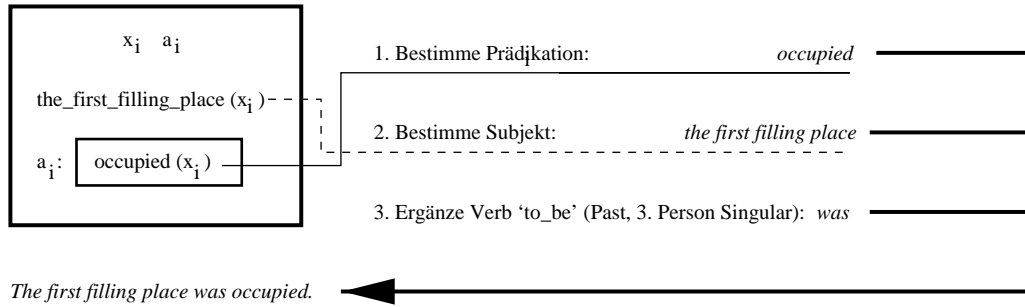


Abbildung 4.27: Satzbauregel zur Transformation einer Prädikation in einen natürlichsprachlichen Satz.

sein (hier: Fahrzeuge, also Neutrum). Zum anderen muß zu der indefiniten Beschreibung des Kernindividuums der passende Artikel bestimmt werden (unbestimmter Artikel im Singular: 'a'). Dabei ist im Beispiel der Sonderfall zu berücksichtigen, daß der Artikel und die indefinite Beschreibung ('other car') verschmelzen zu der Phrase 'another car'.

#### 4.5.5 Prädikation

Abbildung 4.27 demonstriert die Transformation einer Prädikation in einen natürlichsprachlichen Satz. Im Gegensatz zu einer Ablaufbeschreibung mit Bewegungsverb kann eine Prädikation zwar ebenfalls eine Verbform enthalten (hier: *occupied*), diese wird jedoch adjektivisch gebraucht. Deutlicher lassen sich Prädikation und Ablaufbeschreibung mit Bewegungsverb durch die Stellung der DRen in der Ablaufbeschreibung unterscheiden. Während in der Ablaufbeschreibung der DR des Agens (z.B. DR  $x$  in Abbildung 4.25) vor der Verbform steht, wird der DR einer Prädikation in runden Klammern hinten angestellt. Das Subjekt eines prädikativen Satzes ergibt sich aus der Individuenbezeichnung des zugehörigen DRen (in Abbildung 4.27: DR  $x_i$  mit definierter Beschreibung '*the\_first\_filling\_place*'). Als Verb wird 'to\_be' ergänzt. Die Verbform ergibt sich aus Zahl und Geschlecht der Individuenbezeichnung.

#### 4.5.6 Duplexbedingungen, 1. Teil

Enthält die Ablaufbeschreibung eine Duplexbedingung, so wird beim Satzbau wie in Abbildung 4.28 gezeigt verfahren. Die Nominalphrase setzt sich zusammen aus dem natürlichsprachlichen Quantor, der sich im Mittelteil (Raute) der Duplexbedingung befindet. Die Raute enthält zudem denjenigen DR, über den quantorisiert wird (hier:  $x_i$ ). Dieser verweist bei einer Duplexbedingung stets auf eine indefinite Beschreibung, der der natürlichsprachliche Quantor vorangestellt wird. Diese indefinite Beschreibung befindet sich stets in der linken DRS der Duplexbedingung. Die Verbalphrase wird nach den bereits bekannten Regeln erzeugt (hier: Prädikation). Sie ergibt sich aus der rechten DRS der Duplexbedingung.

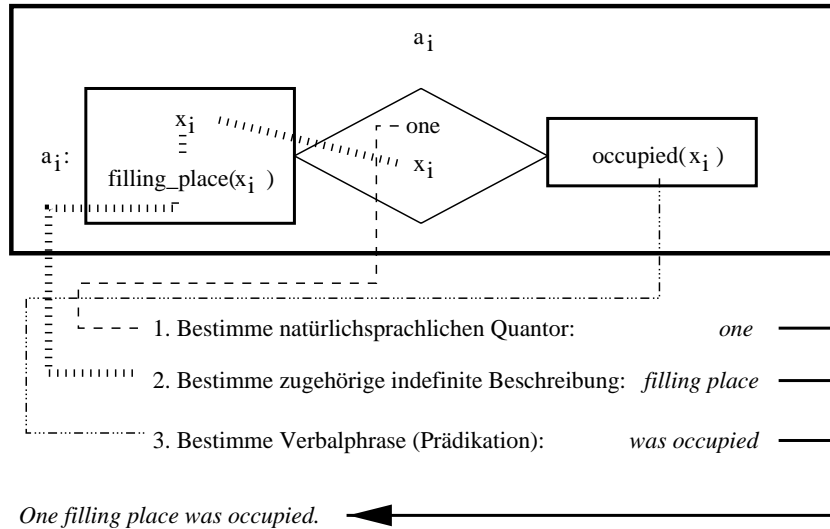


Abbildung 4.28: Satzbauregel für Duplexbedingungen, hier am Beispiel einer Prädikation.

#### 4.5.7 Räumliche Beschreibungsergänzungen

DRS-Bedingungen mit räumlichen Beschreibungsergänzungen enthalten im Relations- teil eine Präposition, die die räumliche Relation angibt, und einen DR, der auf das Referenzobjekt verweist. Als Argument enthält eine solche DRS-Bedingung den DRen, der auf das zu lokalisierende Objekt verweist. Abbildung 4.29 demonstriert die Um- setzung einer räumlichen Beschreibungsergänzung am Beispiel einer Duplexbedingung mit Prädikation. Räumliche Beschreibungsergänzungen werden zur Vereinfachung der Satzbauvorschriften grundsätzlich an den Beginn eines Satzes gestellt und mit einem Komma vom Rest des Satzes abgehoben. Im Beispiel werden die zu lokalisierenden Ob- jekte (Tankplätze) quantorisiert und demzufolge indefinit beschrieben. Das Referenzob- jekt (die obere Tankspur) wird definit beschrieben. Die Zuordnung von lokalisierendem Objekt und Referenzobjekt ergibt sich wie in der Abbildung angegeben.

#### 4.5.8 Duplexbedingungen, 2. Teil

Abbildung 4.30 demonstriert die Behandlung einer Duplexbedingung ohne Prädikation, aber dafür mit Verweis auf eine Ablaufbeschreibung in einer inneren DRS. Die Transfor- mation einer solchen Duplexbedingung in einen natürlichsprachlichen Satz unterscheidet sich von einer Duplexbedingung mit Prädikation in der Ermittlung der Verbalphra- se. Hierfür ist in beiden Fällen die rechte DRS der Duplexbedingung ausschlaggebend. Während sich im Falle einer Prädikation dort das entsprechende Prädikat befindet (Abbildung 4.28), enthält im aktuellen Fall die innere DRS eine Mengenzuordnung. Der DR  $X_i$  verweist auf eine innere DRS, die die Ablaufbeschreibung  $a$  enthält, die der Quantorisierung zugrunde liegt. Die Transformation einer solchen DRS in einen

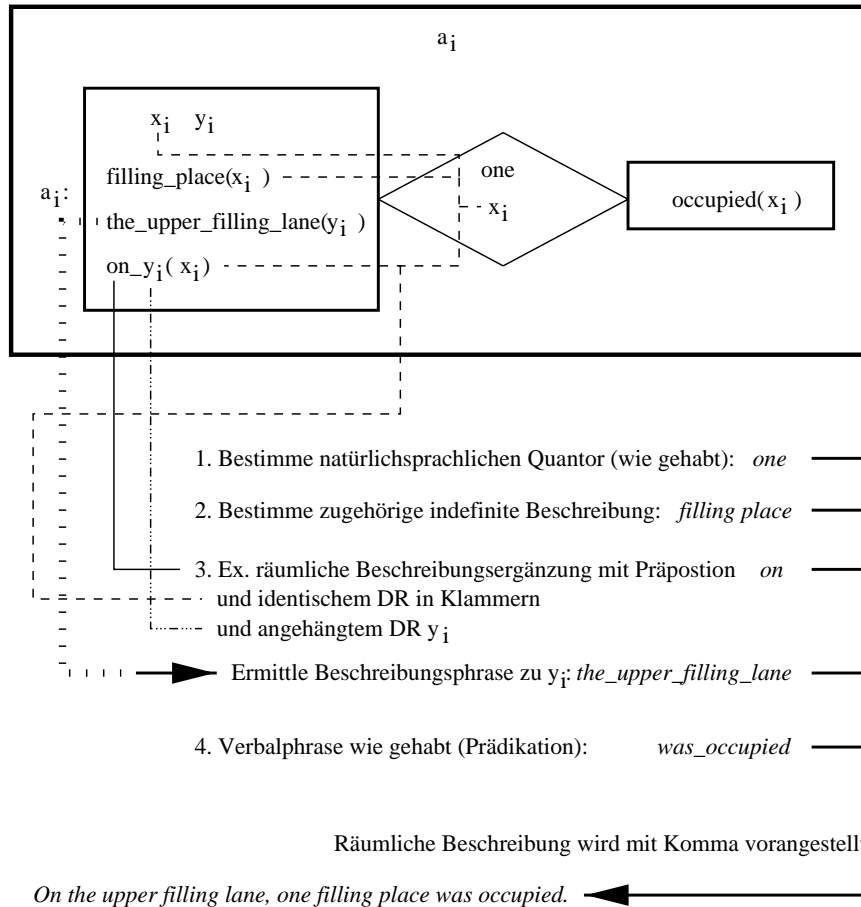


Abbildung 4.29: Satzbauregel für räumliche Beschreibungsergänzungen, hier am Beispiel einer Duplexbedingung mit Prädikation.

Satz ist bereits bekannt. Da die DRS im aktuellen Fall mit einer Duplexbedingung zusammenhängt, werden aus der Ablaufbeschreibung  $a$  nur Verb und Bezugsobjekt versprachlicht. Das Subjekt ergibt sich wie gehabt aus der Raute und der linken DRS der Duplexbedingung.

#### 4.5.9 Satzfolgen basierend auf Duplexbedingungen

Abbildung 4.31 enthält eine DRS mit zwei Ablaufbeschreibungen mit Duplexbedingungen. Die Versprachlichtung des ersten Ablaufs  $a_i$  wurde bereits im vorangegangenen Abschnitt thematisiert. Sie resultiert in dem Satz: *Most vehicles came from Karl-Wilhelm-Straße*. Sei der Ablauf  $a_j$  der im Zeitgraph nachfolgende Ablauf zu  $a_i$ . Dessen Versprachlichtung könnte durch Modifikation der Satzbauregel zur Versprachlichtung von Duplexbedingungen etwa wie folgt aussehen: *Seven of the vehicles coming from Karl-Wilhelm-Straße drove straight ahead*. Bei dieser Art der Versprachlichtung von  $a_j$  bleibt der bestehende inhaltliche Zusammenhang zwischen den Abläufen  $a_i$  und  $a_j$  unberück-



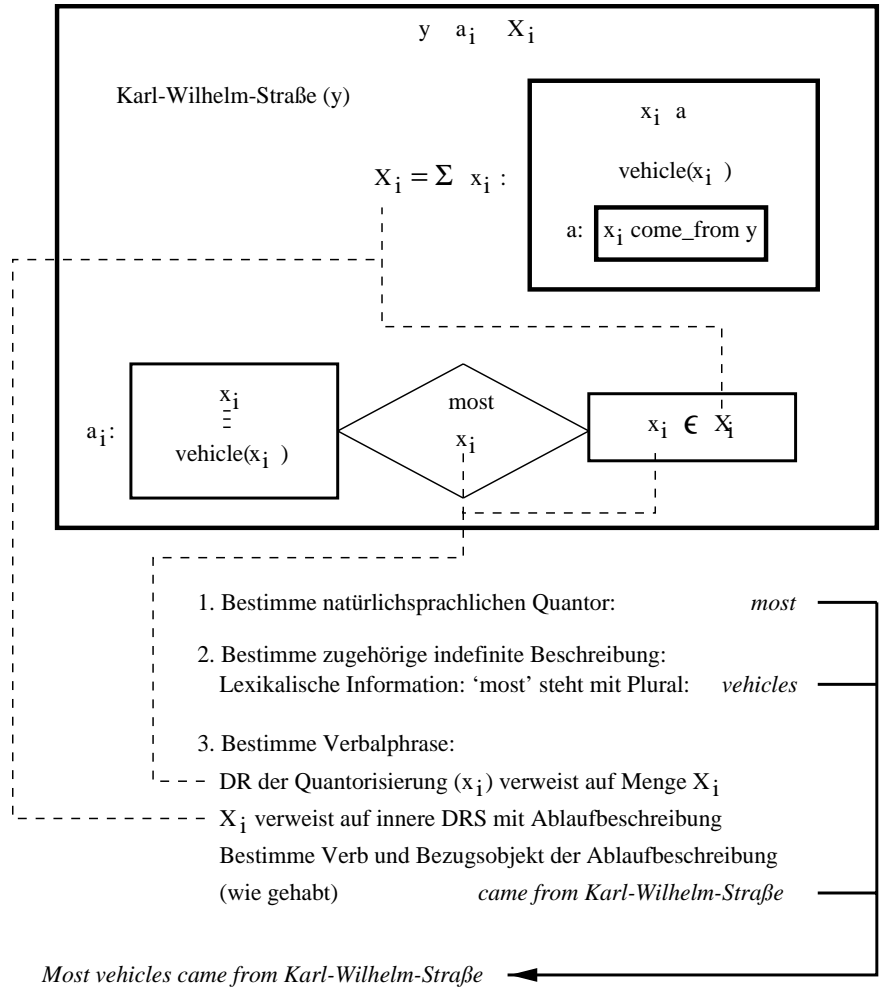


Abbildung 4.30: Satzbauregel für Duplexbedingungen ohne Prädikation.

sichtig: Die Untermenge der Quantorisierung aus  $a_i$  ist identisch mit der Grundmenge der Quantorisierung aus  $a_j$ . Dies zeigt sich darin, daß die DRS-Bedingung der rechten DRS bezüglich  $a_i$  identisch ist mit der DRS-Bedingung der linken DRS bezüglich  $a_j$ . Dieser Sachverhalt kann dazu genutzt werden, um die Formulierung 'the vehicles coming from Karl-Wilhelmstraße' zu ersetzen durch das Pronomen 'them'. Dieses Pronomen verweist dabei eindeutig auf die im Satz davor beschriebene Fahrzeugmenge.

#### 4.5.10 Pronominalisierung

Gegeben sei die in Abbildung 4.32 gezeigte DRS. Die Versprachlichung der Abläufe  $a_1$  und  $a_2$  würde nach den zuvor diskutierten Satzbauregeln folgenden Text ergeben:

(15.1) *Object<sub>4</sub> entered the filling station. Object<sub>4</sub> drove to the first filling place.*

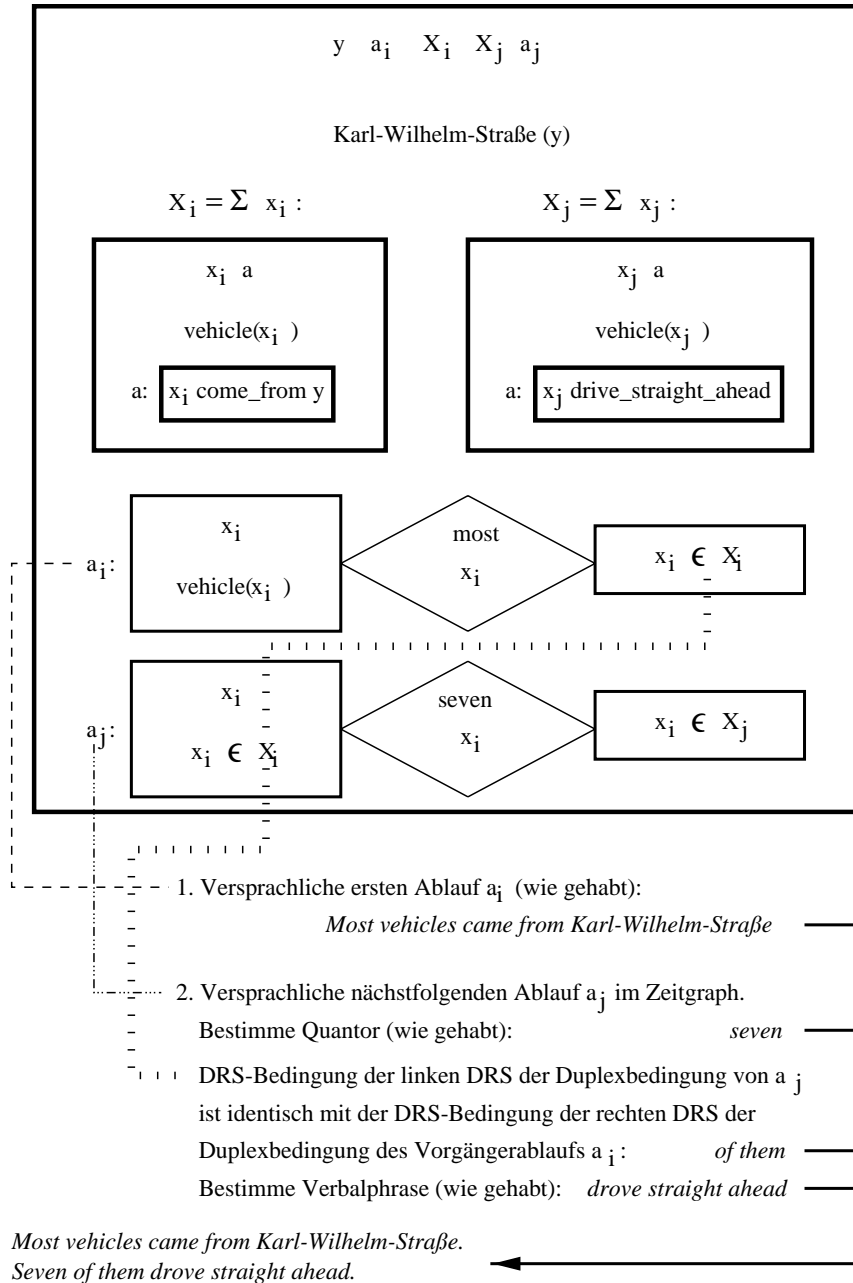


Abbildung 4.31: Satzbauregel für aufeinanderfolgende Ablaufbeschreibungen mit Duplexbedingungen.

Die DRS-Bedingung  $a_1 < a_2$  macht deutlich, daß der erste Ablauf abgeschlossen ist, als der zweite beginnt, denn das Auffahren auf das Tankstellengelände wird an der Tankstelleneinfahrt beobachtet, während das anvisierte Ziel sich für den Beobachter gewöhnlich erst durch das Beobachten des Fahrverhaltens des Agens auf dem Tankstellengelände ausmachen läßt. (15.1) bringt das zeitliche Aufeinanderfolgen der beiden

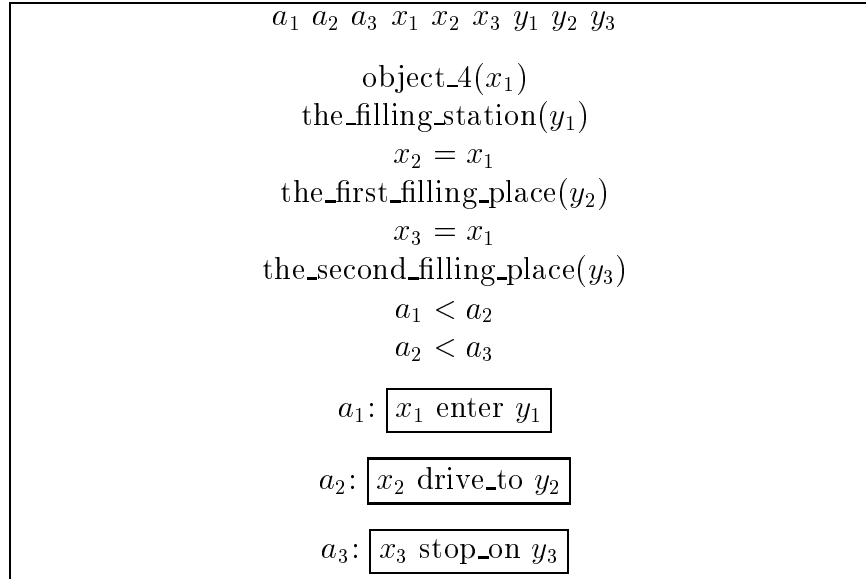


Abbildung 4.32: Beispiel-DRS.

Abläufe dadurch zum Ausdruck, daß der frühere Ablauf im ersten Satz, der darauffolgende Ablauf im zweiten Satz beschrieben wird. Deutlicher könnten beide Abläufe im Text durch den Gebrauch von Temporaladverbien voneinander abgegrenzt werden, so z.B. wenn der zweite Satz durch das Adverb ‘*then*’ eingeleitet werden würde. Störender erweist sich jedoch ein anderes Defizit aus (15.1), nämlich die explizite Erwähnung des Eigennamens in beiden Sätzen. Die damit verbundene Problematik soll anhand zweier weiterer Beispiele motiviert werden:

(15.2) *Object\_4 entered the filling station. Object\_6 drove to the first filling place.*

(15.3) *Object\_4 entered the filling station. It drove to the first filling place.*

(15.2) unterscheidet sich von (15.1) dadurch, daß von zwei verschiedenen Fahrzeugen berichtet wird. (15.2) enthält keine expliziten Hinweise (z.B. durch das Auftreten von Temporaladverbien) darüber, wie die in den beiden Sätzen beschriebenen Abläufe zeitlich miteinander zusammenhängen. Doch die Formulierung von (15.2) legt intuitiv den Schluß nahe, daß beide Abläufe gleichzeitig stattgefunden haben dürften, also Ergebnis einer zusammenhängenden Beobachtung sind. Im Gegensatz dazu drängt sich in (15.3), wo in beiden Sätzen vom selben Fahrzeug die Rede ist, die Vermutung auf, daß die beiden beschriebenen Abläufe nacheinander erfolgten. (15.3) unterscheidet sich von (15.1) allein dadurch, daß im zweiten Satz statt des Eigennamens des Agens als Subjekt das Pronomen ‘*it*’ auftritt, welches auf das Subjekt des ersten Satzes (*Object\_4*) verweist. Zwar könnte algorithmisch das Pronomen ‘*it*’ auch der Tankstelle (*filling\_station*) des ersten Satzes zugerechnet werden, diese Variante kann der (menschliche) Empfänger der Beschreibung (15.3) durch zweifellos vorhandenes Weltwissen ausschließen: Tankstellen

sind unbeweglich und fahren deshalb nicht. Die Diskurse (15.1) und (15.3) sind folglich bedeutungsgleich. Doch (15.3) enthält die vorzuziehende Formulierung des dargestellten Sachverhalts. Dies liegt nicht nur daran, daß die beidermalige explizite Erwähnung des Agens in (15.1) stilistisch unüblich erscheint, sondern auch daran, daß durch die Verwendung des Pronomens in (15.3) beide beschriebenen Abläufe besser voneinander abgegrenzt werden und die Nachfolgebeziehung zwischen den beiden Abläufen implizit besser in Erscheinung tritt. Daraus ergibt sich das Ziel, die Ablaufbeschreibungen  $a_1$  und  $a_2$  nicht mehr in der Form von (15.1), sondern in der Form von (15.3) zu versprachlichen. Dies soll durch folgende Satzbauregel erreicht werden:

*(V1) Gegeben seien zwei im Zeitgraph benachbarte Ablaufbeschreibungen  $a_i$  und  $a_j$  in einer DRS. (V2) Die DRS enthalte zudem die DRS-Bedingung  $a_i < a_j$ . Seien  $x_i$  und  $x_j$  die DRen der Agenten der Abläufe  $a_i$  bzw.  $a_j$ . (V3) Es existiere ferner eine DRS-Bedingung, die den DR  $x_i$  einer Individuenbezeichnung zuführe (in der DRS aus Abbildung 4.32 ist dies die Bedingung  $\text{object}_{\perp 4}(x_1)$ ). (V4) Außerdem existiere die DRS-Bedingung  $x_i = x_j$ . Dann wird bei der Versprachlichung von  $a_j$  statt der expliziten Erwähnung des Agens als Subjekt das zutreffende Personalpronomen verwendet, welches sich aus dem Geschlecht der durch die DRen  $x_j$  bzw.  $x_i$  repräsentierten Agens ergibt.*

Die in der Satzbauregel angegebene Voraussetzung ‘im Zeitgraph benachbart’ stellt sicher, daß die beiden Abläufe  $a_i$  und  $a_j$  im Text in unmittelbar aufeinanderfolgenden Sätzen versprachlicht werden. Durch die DRS-Bedingung  $x_i = x_j$  wird offensichtlich, daß beide Abläufe denselben Agenten besitzen und somit bei der Versprachlichung im zweiten Satz ein entsprechendes Pronomen als Subjekt verwendet werden kann.

Man betrachte noch einmal die DRS aus Abbildung 4.32. Der Ablauf  $a_3$  wurde in den vorstehenden Betrachtungen bisher noch nicht berücksichtigt. Die natürlichsprachliche Beschreibung aller drei Abläufe sollte folgendes Aussehen haben:

(15.4) *Object<sub>⊥4</sub> entered the filling station. It drove to the first filling place. It stopped on the second filling place.*

Durch die oben angegebene Satzbauregel läßt sich noch nicht erreichen, daß  $a_3$  so beschrieben wird, wie es im letzten Satz von (15.4) vorgesehen ist. Der Grund dafür ist, daß keine DRS-Bedingung  $x_2 = x_3$  existiert. Deswegen wird die obige Satzbauregel wie folgt ergänzt:

*Alternativ zu den Voraussetzungen V3 und V4 können folgende Voraussetzungen gelten: (V5) Es existiert ein DR  $x_k$  mit zugehörigen DRS-Bedingungen  $x_i = x_k$  und  $x_j = x_k$ . (V6) Es existiere ferner eine DRS-Bedingung, die den DR  $x_k$  einer Individuenbezeichnung zuführt. Conclusio wie oben.*

### 4.5.11 Weitere Satzbautechniken

Wie schon im vorigen Abschnitt angedeutet wurde, lassen sich automatisch generierte natürlichsprachliche Beschreibungen durch die Einflechtung von **Temporaladverbien** stilistisch verbessern. So zeigt

(16.1) *Object<sub>4</sub> entered the filling station. Then it drove to the first filling place.*

noch deutlicher als (15.3), daß bei Eintritt des zweiten beschriebenen Ablaufs der erste bereits abgeschlossen war. Das Temporaladverb ‘*then*’ scheint also geeignet zu sein Abläufe in der Form “*a<sub>i</sub>. Then a<sub>j</sub>.*” zu versprachlichen, wenn die DRS-Bedingung  $a_i < a_j$  vorliegt. Aber hierfür ebenso geeignet erscheinen auf den ersten Blick die Temporaladverbien *after that*, (*directly*) *afterwards*, *thereupon* oder *later*, um nur einige zu nennen. Diese Temporaladverbien unterscheiden sich jedoch untereinander in semantischen Feinheiten, so daß nicht in jedem Fall zweier Abläufe  $a_i$ ,  $a_j$  mit  $a_i < a_j$  jedes dieser Temporaladverb gleich gut anwendbar ist. So bestimmt beispielsweise der zeitliche Zwischenraum zwischen zwei Abläufen, ob das Temporaladverb ‘*then*’ oder besser das Temporaladverb ‘*later*’ zur zeitlichen Abgrenzung verwendet wird. Bei der Verwendung von ‘*later*’ scheint der zeitliche Zwischenraum größer zu sein (etwa: nach einer gewissen Pause) als bei der Verwendung von ‘*then*’, welches eher ein ‘unmittelbares Aufeinanderfolgen’ impliziert. Doch wie groß ein zeitlicher Zwischenraum sein muß, um zwischen besagten Temporaladverbien trennen zu können, läßt sich nicht kontextunabhängig bestimmen, wie folgende beiden Beispiele zeigen.

(16.2) *A meteor hit the earth. Then the dinosaurs died.*

(16.3) *Object<sub>4</sub> drove on Kriegsstraße. Later, a pedestrian crossed the street.*

In (16.2) wird ‘*then*’ verwendet, was darauf schließen läßt, daß nur kurze Zeit zwischen dem Einschlag des Meteors und dem Sterben des letzten Dinosauriers verging. Geht man davon aus, daß tatsächlich ein Meteoriteneinschlag für das Aussterben der Dinosaurier verantwortlich war, dann dürfte besagte ‘kurze Zeit’ zumindest im Bereich von Monaten oder Jahren anzusiedeln sein, keinesfalls aber im Bereich von Sekunden oder Minuten (den Dinosauriern ist der Meteor schließlich nicht direkt auf den Kopf gefallen, zumindest nicht allen, sondern sie dürften an den klimatischen Folgen des Einschlags zugrunde gegangen sein). Der zeitliche Zwischenraum in (16.3) dürfte gefühlsmäßig maximal im Stundenbereich anzusiedeln sein, ist also viel kürzer als in (16.2), obwohl in (16.3) statt dem ‘kurzintervalligen *then*’ das langintervallige ‘*later*’ verwendet wurde. Zu beachten ist außerdem, daß sich bisher aus DRSen die ‘Pausendauer’ zwischen zwei Abläufen gar nicht bestimmen läßt, denn ‘ $<$ ’ ist die einzige bestehende temporale Relation für aufeinanderfolgende Abläufe, und diese liefert nur qualitative, aber keine quantitative Information.

Bei der Assoziation von Temporaladverbien zur Beschreibung des temporalen Zusammenhangs zweier Abläufe ist außerdem zu beachten, daß Temporaladverbien oft nicht

nur eine temporale, sondern auch eine kausale Beziehung zwischen Abläufen herstellen. So wird in (16.2) implizit durch die Verwendung von ‘*then*’ auch ein *kausaler* Zusammenhang zwischen dem Meteoriteneinschlag und dem Aussterben der Dinosaurier hergestellt. Dieser kausale Zusammenhang wäre nicht so offensichtlich betont, wenn statt ‘*then*’ das Temporaladverb ‘*later*’ verwendet worden wäre.

Als Fazit der voranstehenden Überlegungen läßt sich feststellen, daß die Einflechtung ‘passender’ temporaler Beschreibungszusätze in einen Text nicht allein aufgrund der jeweils gültigen temporalen Relation getroffen werden kann, sondern von verschiedenartigsten semantischen oder kontextuellen Faktoren abhängig ist. Dies zeigt sich nicht nur für Relationen der Form  $a_i < a_j$ , sondern genauso oder noch mehr für zeitliche Überlappungen ( $a_i \circ a_j$ ). Gegeben seien folgende Fakten:

(16.4) 100 : 200 ! enter(object\_4, the\_filling\_station, 1).

(16.5) 100 : 200 ! turn\_on(object\_4, the\_lights, 2).

Ordnet man den beiden Abläufen die DRSen  $a_1$  bzw.  $a_2$  zu, so resultiert die DRS-Bedingung  $a_1 \circ a_2$ . Man könnte geneigt sein, (16.4,5) durch (16.6) natürlichsprachlich zu formulieren.

(16.6) *Object\_4 entered the filling station. At the same time, it turned its lights on.*

Die Assoziation der adverbialen Bestimmung ‘*at the same time*’ zu der DRS-Bedingung  $a_1 \circ a_2$  führt jedoch im Falle von (16.7,8) zu einer stilistisch mäßigen Beschreibung.

(16.7) 100 : 200 ! want\_to\_overtake(object\_4, object\_6, 1).

(16.8) 100 : 200 ! change\_to(object\_4, the\_passing\_lane, 2).

Hier erscheint es angebrachter, den kausalen Zusammenhang zwischen den beiden Fakten zu betonen, statt den temporalen. Somit ist die Formulierung aus (16.9) derjenigen aus (16.10) vorzuziehen.

(16.9) *Object\_4 wanted to overtake object\_6. It, therefore, changed to the passing lane.*

(16.10) *Object\_4 wanted to overtake object\_6. At the same time, it changed to the passing lane.*

Um die Diskussion auf die Spitze zu treiben sei abschließend ein Beispiel angegeben, aus dem sich die Verwendung des Temporaladverbs ‘*then*’ anbietet, obwohl als zugehörige DRS-Bedingung ‘ $a_1 \circ a_2$ ’ extrahiert wurde. Man betrachte folgende Fakten.

(16.11) 100 : 180 ! turn\_left\_into(object\_4, Kriegsstraße, 1).

(16.12) 120 : 200 ! turn\_left\_into(object\_6, Kriegsstraße, 2).

Gemeint sei, daß sich beide Fahrzeuge auf einer Linksabbiegespur zur Kriegsstraße befindet, wobei Objekt\_6 Objekt\_4 folgt. Obwohl bei der Analyse der Gültigkeitsintervalle (16.11,12) als DRS-Bedingung ‘ $a_1 \circ a_2$ ’ extrahiert werden würde (schließlich liegt eine zeitliche Überlappung der beiden Abläufe vor), stellt (16.13) eine zumindest mögliche natürlichsprachliche Beschreibung dar.

(16.13) *Object\_4 turned left into Kriegsstraße. Then object\_6 turned left into Kriegsstraße.*

Als endgültiges Fazit dieser Betrachtungen sei festgestellt, daß der berechtigte Wunsch, automatisch generierte natürlichsprachliche Beschreibungen durch die gezielte Einflechtung temporaler Beschreibungszusätze stilistisch aufwerten zu wollen, fundierte linguistische Vorbetrachtungen erfordert. Im Rahmen der vorliegenden Arbeit werden die entsprechenden Fragestellungen nicht annähernd gelöst. Temporale Beschreibungszusätze werden bei der Textgenerierung in überschaubaren Einzelfällen dennoch verwendet. So wird beispielsweise das Temporaladverb ‘*then*’ assoziiert, wenn zwei aufeinanderfolgende Abläufe versprachlicht werden sollen, die erstens beide durch ein Bewegungsverb beschrieben werden, und die zweitens von ein und demselben Agenten durchgeführt werden. Handelt es sich um zwei verschiedene Agenten, wird stattdessen ‘*later*’ verwendet. ‘*In the meantime*’ wird bei zeitlicher Überlappung verwendet, wenn beide Abläufe durch ein Bewegungsverb beschrieben werden und verschiedene Agenten besitzen. Diese Festlegungen sind jedoch rein provisorischer Natur.

Neben temporalen sind auch **kausale Beschreibungszusätze** für die Textgenerierung interessant. Bei den Ausführungen über die DRS-Generierung aus begrifflichen Beschreibungen wurde bereits die Notwendigkeit der Repräsentation von kausalen Beziehungen zwischen Abläufen betont und andiskutiert. Daraus erwächst der Wunsch, eine solchermaßen repräsentierte Kausalbeziehung durch Verwendung von Kausaladverbien wie *because of*, *because*, *in order to* und ähnliche explizit in den zu generierenden Text einzuflechten.

Doch mit der Einflechtung kausaler und temporaler Beschreibungszusätze allein sind die stilistischen Textgestaltungsmöglichkeiten längst nicht ausgeschöpft. Man betrachte folgendes Textbeispiel:

(16.14) *Object\_4 entered the filling station, drove to the first filling place, and finally stopped on the second one.*

(16.14) enthält dieselbe inhaltliche Information wie (15.4), ist aber stilistisch besser formuliert. Als zusätzliche Stilmittel wurden die Reihung aufeinanderfolgender, sich gegenseitig ablösender Abläufe in einem Satz, die Verwendung sprachlicher Ellipsen (‘*one*’) und adverbiale Ergänzungen (‘*finally*’) verwendet. Darüber hinaus sind noch viele andere zusätzliche Stilmittel denkbar (z.B. die explizite Kenntlichmachung von Wiederholungen von Abläufen mittels ‘*again*’ oder die explizite Erwähnung von Analogien im Verhalten verschiedener Agenten durch ‘*as well*’ etc.).

Der DRS-Formalismus erscheint grundsätzlich mächtig genug zu sein, um die Basis zu schaffen für die Generierung natürlichsprachlicher Texte, welche die Einflechtung verschiedenster stilistischer Beschreibungszusätze erlauben. Das Problem liegt jedoch nicht in der Beantwortung der Frage, wie man bestimmte stilistische Beschreibungszusätze aus einer DRS ableiten kann. Vielmehr stellt sich das Problem, den Entscheidungsprozeß zu formulieren, **wann** die Einflechtung eines bestimmten Beschreibungszusatz überhaupt sinnvoll ist. Eine Reihung von Abläufen wie in (16.14) beispielsweise wirkt vielleicht stilvoll, enthält aber ein Text mehrere aufeinanderfolgende Sätze, die im selben Stil aufgebaut sind, so wirkt dies auf den Leser schnell langweilig oder störend. Auch hier zeigt sich, daß die Generierung einer stilvollen natürlichsprachlichen Beschreibung sich nicht auf die grundsätzliche Fähigkeit zur Einflechtung bestimmter Stilmittel beschränkt, sondern in erster Linie in deren wohl dosiertem Einsatz begründet ist. Ein gewisses ‘Sprachgefühl’ erscheint erforderlich, welches sich bisher noch nicht ausreichend in ein automatisches System umsetzen läßt.

## 4.6 DRS-Manipulation

In den vorausgegangenen zwei Abschnitten wurde einerseits die Transformation einer begrifflichen Beschreibung in eine DRS diskutiert, andererseits die Transformation einer DRS in einen natürlichsprachlichen Text. Noch unbeantwortet ist die Frage, wie die DRS selbst als Plattform für logikbasierte Manipulation genutzt werden kann. Es steht also die Installierung eines Schlußfolgerungsmechanismus im Raum, welcher im Sinne einer Abbildung eine DRS in eine andere DRS überführt. Der Zweck einer solchen Abbildung könnte u.a. in der Eliminierung von Redundanzen in der Ursprungs-DRS, in der Ableitung zusätzlicher Inhalte aus zusätzlich einfließendem Weltwissen oder in der Erzielung einer höheren Abstraktionsstufe liegen. Anwendungsabhängig lassen sich drei realistische Szenarien für den Nutzen logikbasierter DRS-Manipulation vorstellen:

1. Ein Anwender stellt eine natürlichsprachliche Anfrage an das System. **Anhand dieser Anfrage wird aus einer bestehenden DRS** die Antwort abgeleitet. Der Schlußfolgerungsmechanismus leitet aus der bestehenden DRS eine ‘kleinere’ DRS ab, die nur diejenige Information enthält, die zur Beantwortung der Anfrage relevant ist (**Antwort-DRS**). Aus dieser neuen, pragmatischen DRS wird dann der Antworttext **abgeleitet**.
2. Die in einer systemintern vorhandenen DRS enthaltene Information soll versprachlicht werden. Die DRS sei nicht pragmatisch, d.h. sie enthält Redundanzen und entbehrt einer textgestalterischen Ordnung. Demzufolge soll der Schlußfolgerungsmechanismus die **vorhandene, nicht pragmatische DRS in eine pragmatische DRS überführen**, aus der unmittelbar ein stilistisch ansprechender natürlichsprachlicher Text abgeleitet werden kann.
3. Die DRS soll als **Basis für erweiterte Schlußfolgerungen** genutzt werden.



Ein entsprechender Schlußfolgerungsmechanismus ist bisher nicht vorhanden. Im folgenden sind einige eigene strategische Maßnahmen und Überlegungen aufgezeichnet, die die oben angegebenen Systemanforderungen zu bewerkstelligen oder zu überbrücken helfen.

Zur **Generierung einer Antwort-DRS zu einer Anfrage** (Punkt Eins der obigen Aufzählung) ist folgendes zu sagen: Zum jetzigen Stand der Entwicklung kann der Benutzer noch keine gezielte natürlichsprachliche Anfrage an das System stellen. Er kann aber statt dessen ein Erzählziel formulieren. Anhand dieses Erzählziels kann jedoch bisher noch nicht aus einer bereits bestehenden DRS eine Antwort abgeleitet werden. Statt dessen wird erzählzielabhängig die begriffliche Auswertung angestoßen, welche eine begriffliche Beschreibung extrahiert, aus der dann unmittelbar eine pragmatische DRS erstellt wird. Aus dieser DRS wird ein natürlichsprachlicher Text abgeleitet, der das angegebene Erzählziel erfüllt. Auf diese Weise wird derzeit die DRS-Manipulation zur Generierung einer Antwort-DRS 'eingespart'. Für die Zukunft erscheint es jedoch wünschenswert, nicht zuletzt aus Gründen der Laufzeitoptimierung, entsprechende Möglichkeiten zur DRS-Manipulation zu schaffen. Dies hätte den Vorteil, daß nicht bei jeder neuen Anfrage die begriffliche Auswertung neu angestoßen werden müßte, um eine neue DRS zu erzeugen, auch wenn vielleicht schon als Ergebnis früherer Anfragen eine DRS vorhanden ist, die die gesuchte Information bereits enthält, und man schlicht nicht in der Lage ist, dies zu erkennen oder die Information entsprechend zu extrahieren.

Das Problem der **Ableitung einer pragmatischen DRS aus einer nicht pragmatischen DRS** (Punkt Zwei der obigen Aufzählung) stellt sich derzeit nicht. Die erzählzielbasierte begriffliche Auswertung liefert generell nur pragmatische DRSen als Ergebnis. Wurde daraus der dem Erzählziel entsprechende natürlichsprachliche Text generiert, wird in der Folge die DRS wieder gelöscht und mit dem nächsten zu bewältigenden Erzählziel eine neue, pragmatische DRS berechnet. Die Effizienz der Textgenerierung ließe sich dabei dann erhöhen, wenn bereits extrahierte Information in einer DRS dauerhaft gespeichert und im Bedarfsfall ohne Neuberechnung in eine neue DRS eingehen kann. Man kann jedoch nicht davon ausgehen, daß bei der Vereinigung zweier pragmatischer DRSen automatisch eine pragmatische DRS entsteht. Im Gegenteil wird man aus der vereinigten DRS eine pragmatische erst ableiten müssen. Dies wird bisher jedoch noch nicht realisiert.

Der obige dritte Punkt sieht die Realisierung eines Schlußfolgerungsmechanismus zur Ableitung zusätzlicher DRS-Bedingungen aus einer bestehenden DRS vor. Zum Hintergrund für die zusätzliche Abstraktionsstufe, die durch diese Art der DRS-Manipulation eingenommen würde, beachte man, daß bei der Transformation einer begrifflichen Beschreibung in eine DRS zum jetzigen Stand der Entwicklung stets eine eigene, neue DRS aufgebaut wird. Anders ausgedrückt bedeutet dies, daß die Ursprungs-DRS, in die eine neue begriffliche Beschreibung transformiert wird, stets leer ist. Es stellt sich die Frage, ob man neu berechnete begriffliche Beschreibungen nicht besser in DRSen transformiert, die nicht leer sind, das heißt, bereits bestimmte Zusatzinformationen re-

präsentieren. Diese Zusatzinformation könnte zum einen in zusätzlichem Wissen über den Diskursbereich, zum anderen in linguistischem Wissen über die Ableitung von Text aus DRSen bestehen. Prinzipiell könnte das gesamte Weltwissen, das bisher in den Situationsgraphen repräsentiert wird, auch in Form einer DRS repräsentiert werden. Die DRS könnte dann das Schema bilden, das anhand berechneter Bildauswertungsdaten ausgeprägt wird. Diese Überlegungen wurden erstmals in [Nagel 99] angestellt. Der Autor regt an, das ganze Diskursbereichswissen natürlichsprachlich zu formulieren. Dieser Text könnte anhand von [Kamp&Reyle 93] in eine DRS überführt werden. Diese DRS bildet dann das Schema für eine begriffliche Auswertung. Für den Systemverwalter würde diese Vorgehensweise eine erhebliche Arbeitserleichterung bedeuten. An die Stelle des Entwurfs eines komplexen Situationsgraphen zur schematischen Repräsentation von Diskursbereichswissen genügt die Formulierung eines entsprechenden natürlichsprachlichen Textes, der lediglich sprachsyntaktische Anforderungen erfüllen und die entsprechenden Inhalte enthalten müßte.

## 4.7 Lexika

[Kamp&Rossdeutscher 94a] beschäftigen sich mit der Fragestellung, wie ein Lexikon aufgebaut sein muß, um die systematische Repräsentation von Texten nach Maßgabe der eigenen DRT geeignet zu unterstützen. Die Überlegungen in dieser Arbeit konzentrieren sich auf den transitiven wie intransitiven Gebrauch des deutschen Verbs ‘heilen’. Dabei werden mögliche Argumentenstrukturen dieses Verbs und deren ‘thematische Rolle’ betrachtet. Als *thematische Rolle* werden funktionale Relationen zwischen bestimmten Abläufen und den Partizipanten dieser Abläufe bezeichnet. Die möglichen thematischen Rollen, die die Argumente eines in Form eines Prädikats beschriebenen Ablaufs einnehmen können, werden als lexikalischer Eintrag zu dem betreffenden Verb benutzt. Für das Beispiel ‘heilen’ wird ein solcher lexikalischer Eintrag konkret entwickelt, wodurch die semantische Repräsentation einiger Sätze, die dieses Verb enthalten, ermöglicht wird. Dadurch wird ein Formalismus geschaffen, der geeignet erscheint, Lexika zu beschreiben, welche nicht nur zweckgebunden alle erforderlichen Informationen enthalten, sondern diese zusätzlich in einer Form repräsentieren, die den Zugriff auf diese Information vereinfacht. Das Lexikon bestimmt entscheidend, ob eine hinreichend differenzierte semantische Repräsentation von Texten möglich wird. Diese Differenziertheit zeigt sich in der Fähigkeit, Inferenzen zu ermöglichen, die aus einem repräsentierten Text gezogen werden sollen. Als solche Inferenzen kommen grundsätzlich folgende in Frage:

- Die automatische Antwortgenerierung zu einer Anfrage;
- die automatische maschinelle Übersetzung von Texten;
- die Erzeugung der semantischen Repräsentation selbst und

- die Rekonstruktion temporaler Bezüge.

[Kamp&Rossdeutscher 94b] demonstrieren die wichtige Rolle, die lexikalische Information bei der Erzeugung einer rechnerinternen Repräsentation spielt, anhand von Beispielen von Ableitungen aus bestehenden Textrepräsentationen. Dabei wird auch im speziellen die Bedeutung des Wortes ‘wieder’ untersucht und für dieses Wort ein lexikalischer Eintrag erzeugt.

Die Konstruktion hinreichender lexikalischer Einträge (beispielsweise für das Wort ‘wieder’) ist deshalb so schwer, da die Autoren sich mit Texten konfrontiert sehen, deren Semantik sie repräsentieren wollen, über deren Inhalt und Struktur sie aber keine Vorannahmen treffen wollen. Um dies hinreichend robust bewältigen zu können, müssen sie idealerweise jede semantische Bedeutung des Wortes ‘wieder’ erfassen können. Vier kleine Beispiele sollen den unterschiedlichen Gebrauch dieses Wortes illustrieren:

(17.1) “*Es regnet wieder.*”

(17.2) “*Es war wieder Möller, der auf das Tor schoß.*”

(17.3) “*Da hat schon wieder einer vor unserer Hofeinfahrt geparkt.*”

(17.4) “*Das Fahrzeug fährt wieder an.*”

In (17.1) weist das Wort ‘wieder’ auf das erneute Auftreten eines Ereignisses hin, das innerhalb einer lokalen zeitlichen Umgebung in vergleichbarer Weise schon mindestens einmal auftrat. Betrachtet man (17.2) als Kommentar zur Übertragung eines Fußballspiels, so wird durch diese Äußerung nicht notwendigerweise besagt, daß das Agens ‘Möller’ schon zuvor auf das Tor schoß (also wie in (17.1) das beschriebene Ereignis in vergleichbarer Weise schon einmal auftrat). Es kann statt dessen ebenso gut sein, daß durch (17.2) lediglich ausgesagt werden soll, daß das Agens ‘Möller’ bereits zuvor auffällig in Erscheinung getreten ist. In (17.3) wird wie in (17.1) die Wiederholung eines bestimmten Ereignisses betont, wobei allerdings im Gegensatz zu (17.2) nicht verlangt wird, daß wieder dasselbe Agens beteiligt gewesen sein muß. In (17.4) wird nicht ausgesagt, daß bereits zuvor schon einmal das Anfahren des Fahrzeugs beobachtet wurde, also die Wiederholung eines bestimmten Ereignisses vorliegt. Statt dessen wird das Wort ‘wieder’ verwendet, um die Rückkehr in einen vormals beobachteten Zustand (‘fahrend’) zu kennzeichnen.

Auch im Rahmen der vorliegenden Arbeit wird zur Ableitung natürlichsprachlicher Beschreibungen durch Bildfolgenauswertung auf Begriffslexika zurückgegriffen. Deren Eintragungen sind jedoch im Vergleich wesentlich simpler. Dies können sie auch sein, da sie zur Textgenerierung und nicht zur Textanalyse verwendet werden. Bei der Übertragung einer in der begrifflichen Schicht berechneten begrifflichen Beschreibung wird auf Lexika zurückgegriffen, die Informationen über den Gebrauch der einzelnen Wörter der begrifflichen Beschreibung enthalten. Je allgemeiner der lexikalische Eintrag, desto beschränkter ist die sprachliche Ausdrucksfähigkeit. Diese Aussage gilt auch für die Ableitung des natürlichsprachlichen Textes aus der DRS, wenn zum zweiten Mal auf

die Lexika zugegriffen wird. Für das Beispiel des Wortes 'wieder' bedeutet dies, daß der Gebrauch des Wortes auf verwandte Fälle zu (17.1) eingeschränkt wird. Dadurch wird das Problem der Herleitung eines vollständigen lexikalischen Eintrags zu Lasten der sprachlichen Ausdrucksmöglichkeit umgangen. Bei einer zukünftigen schrittweisen Steigerung der sprachlichen Ausdrucksfähigkeit des Systems (für die ja diese Arbeit u.a. eine Grundlage geben will) müssen die entsprechenden Lexika sukzessive erweitert werden.

Zum jetzigen Stand der Entwicklung wird für jede relevante Wortart (Adjektiv, Adverb, Artikel, Nomen, Personalpronomen, Präpositionen, Verben etc.) ein eigenes Lexikon verwendet. Im Falle der Personalpronomen wird zu jedem Vertreter Kasus, Person, Zahl und ggf. Geschlecht eingetragen. Das Verblexikon enthält für jeden Eintrag Angaben über die verschiedenen Zeitformen des Verbs (Present Tense, Past Tense und Past Participle) und die Verlaufsform. Die Lexika enthalten ferner alle bisher vertretbaren Vereinfachungen. So wird beispielsweise bei den Nomen auf den Vermerk des Geschlechts der einzelnen Nomina verzichtet. Der Grund: Die auftretenden Agenten sind alle geschlechtslos (Fahrzeuge), mögliche Bezugsobjekte sind ebenfalls stets geschlechtslos (Fahrzeuge, stationäre Objekte oder abstrakte Objekte). Man beachte, daß englischsprachige Texte erzeugt werden, und es sich deshalb um englischsprachige Lexika handelt.

# Kapitel 5

## Ergebnisse, Zusammenfassung und Ausblick

### 5.1 Testszzenarien

Zur Überprüfung der in dieser Arbeit vorgestellten Methoden wurden zahlreiche Experimente durchgeführt. Diese Experimente basieren auf realen Daten. An mehreren Straßenverkehrskreuzungen und an einer Tankstelle wurden mit Hilfe einer stationären Standard-Videokamera Bildfolgen aufgezeichnet, welche durch das am Institut entwickelte Bildfolgenauswertungssystem ausgewertet wurden. Dabei wurden für jedes verfolgte Fahrzeug mehrere, perspektivisch wechselnde, natürlichsprachliche Beschreibungen seines Verhaltens berechnet. Einige Ergebnisse sind im Einzelnen in Anhang A dieser Arbeit dokumentiert.

Abbildung 1.1 zeigt vier Einzelbilder aus einer Straßenverkehrsszene, die an einer Tankstelle aufgezeichnet wurde (Tankstellenszene, [Damm 94]). Die Bildfolge umfaßt ca. 8.000 Halbbilder. Dies entspricht bei einer Abtastrate von 50 Halbbildern pro Sekunde einer zeitlichen Dauer der Bildfolge von ca. drei Minuten. Während dieser Zeit treten in der Szene verschiedene Fahrzeuge auf, die zielgerichtet bestimmte Zapfsäulen ansteuern und später wieder verlassen. Die Fahrmanöver wurden inszeniert und sind daher nicht zufällig. Dabei wurde vereinfachend davon ausgegangen, daß jeder Fahrer an jeder Zapfsäule den für sein Fahrzeug relevanten Kraftstoff erhalten kann. Damit bestimmt nicht der Typ der Zapfsäule, sondern allein der jeweils aktuelle Belegungszustand der Zapfsäulen das Verhalten der Fahrzeuge bzw. deren Fahrer. Nähere Einzelheiten über die der Auswertung zugrunde liegenden Situationsmodellierung entnehme man Abschnitt 3.4.2 oder dem in Anhang B abgedruckten Situationsgraphenbaum.

Zusätzlich zur Tankstellenszene wurden Straßenverkehrsszenen an fünf innerstädtischen Kreuzungen aufgezeichnet. Diese Szenen wurden nicht inszeniert, sondern geben die zufällige Verkehrssituation im Moment der Aufzeichnung wieder. Wie bei der Tankstellenszene betrug die Abtastrate 50 Halbbilder pro Sekunde, aufgezeichnet mit einer

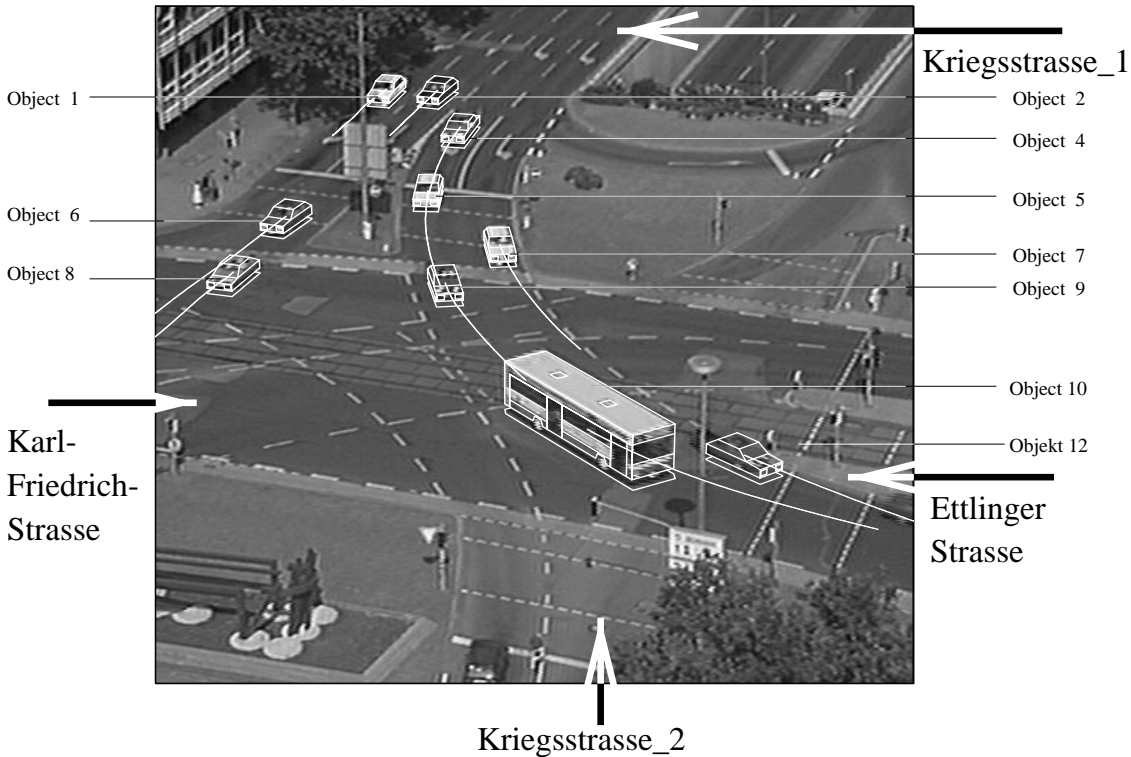


Abbildung 5.1: Straßenverkehrsszene am Ettlinger Tor.

stationären, monokularen Standard-Videokamera. Eine Auswahl an Auswertungsergebnissen ist in Anhang A dieser Arbeit angegeben. Dort sind auch einige Einzelbilder zu den verschiedenen Kreuzungsszenen enthalten. Die Aufnahmedauer der einzelnen Bildfolgen unterscheiden sich.

Die Bildfolgen 'Ettlinger-Tor-Szene' (Abbildung 5.1) und die 'Alte-Durlacher-Tor-Szene' sind älteren Datums und daher sehr kurz: Die Aufnahmedauer liegt, aufgrund der damals beschränkten Speicherkapazität der Rechner, jeweils bei unter zwei Sekunden. Dies hat bei der Berechnung natürlichsprachlicher Beschreibungen den Nachteil, daß die einzelnen Fahrzeuge nicht über den gesamten Zeitraum ihres Aufenthalts auf der jeweiligen Kreuzung erfaßt wurden. Darunter kann die Qualität der generierten natürlichsprachlichen Beschreibung leiden. Ebenfalls älteren Datums, aber doch von fortgeschrittener Länge ist die Nibelungenplatzszene, die dritte ausgewertete Kreuzungsszene. Die 'Rheinhafenszene' unterscheidet sich von allen anderen Kreuzungsszenen dadurch, daß an der beobachteten Kreuzung keine Ampelsteuerung vorliegt, sondern die Vorfahrt durch Verkehrszeichen geregelt ist. Dadurch läßt sich besonders gut das Verhalten von nicht vorfahrtsberechtigten Fahrzeugen analysieren, welche versuchen, in den fließenden Verkehr einzufädeln. Die 'Neue-Durlacher-Tor-Szene' beinhaltet das umfangreichste Datenmaterial. Sie umfaßt insgesamt zwölf Bildfolgen (ausführlich dokumentiert in [Haag&Nagel 99]) über ca. 43.500 Halbbildzeitpunkte (entspricht 870 Sekunden bzw. über vierzehn Minuten). Inzwischen besteht eine Glasfaserverbindung vom Institut für

Algorithmen und Kognitive Systeme zu einer fest installierten Videokamera auf dem Dach des Bauingenieurgebäudes der Universität Karlsruhe, mit der besagte Kreuzung rund um die Uhr beobachtet werden kann. Die Auswertung neuen Datenmaterials wird, nach Kalibrierung dieser Kamera, folglich jederzeit und unmittelbar möglich sein, sofern die Beleuchtungsverhältnisse es gestatten, Fahrzeuge mit hinreichendem Kontrast gegenüber dem stationären Bildvorder- und -hintergrund zu erfassen.

## 5.2 Auswertungsergebnisse

Eine Auswahl an berechneten natürlichsprachlichen Beschreibungen ist in Anhang A dieser Arbeit dokumentiert. In dem vorliegenden Abschnitt werden die einzelnen Auswertungsschritte an einem einfachen Beispiel beschrieben. Es bezieht sich auf das Fahrzeug ‘Object\_10’ der Ettliger-Tor-Szene (Abbildung 5.1). Dieses Fahrzeug wurde im Zeitintervall [5,95] verfolgt. Die signalnahe Auswertung ergibt eine geometrische Beschreibung, die in Ausschnitten wie folgt angegeben ist. Sie enthält Schätzwerte für die (x,y)-Position des Fahrzeugs in der Szene, seiner Orientierung, Geschwindigkeit und Winkelgeschwindigkeit.

Bild#	x[m](+-dx)	y[m](+-dy)	$\theta$ [g](+-d $\theta$ )	v[m/s](+-dv)	w[g/s](+-dw)
5	12.65(0.1310)	-7.60(0.1450)	-33.99(1.340)	9.52(5.00)	0.00(28.6)
6	12.81(0.0974)	-7.71(0.1030)	-34.13(0.977)	9.49(4.59)	-1.07(27.4)
7	12.96(0.0944)	-7.82(0.0911)	-34.15(0.905)	9.49(3.77)	-1.07(24.4)
8	13.12(0.0943)	-7.92(0.0853)	-34.12(0.904)	9.50(2.99)	-0.28(20.6)
9	13.28(0.0915)	-8.01(0.0801)	-34.00(0.895)	9.53(2.41)	1.53(16.9)
10	13.45(0.0883)	-8.11(0.0757)	-33.84(0.874)	9.57(2.04)	3.11(13.8)
11	13.61(0.0853)	-8.21(0.0721)	-33.62(0.850)	9.65(1.82)	4.84(11.5)
12	13.79(0.0819)	-8.31(0.0686)	-33.37(0.815)	9.77(1.68)	6.35(9.63)
13	13.97(0.0779)	-8.41(0.0648)	-33.09(0.772)	9.90(1.59)	7.69(8.12)
14	14.14(0.0753)	-8.50(0.0620)	-32.76(0.738)	9.92(1.54)	8.99(6.99)
15	14.31(0.0740)	-8.60(0.0600)	-32.48(0.710)	9.95(1.53)	9.74(6.12)
16	14.48(0.0733)	-8.69(0.0586)	-32.19(0.688)	9.92(1.52)	10.32(5.43)
17	14.66(0.0730)	-8.78(0.0575)	-31.93(0.668)	9.94(1.51)	10.63(4.87)
18	14.83(0.0722)	-8.88(0.0561)	-31.66(0.646)	9.92(1.51)	10.96(4.38)
19	15.00(0.0717)	-8.98(0.0550)	-31.44(0.625)	9.92(1.50)	10.96(3.98)
20	15.18(0.0714)	-9.07(0.0541)	-31.18(0.608)	9.88(1.50)	11.18(3.65)
21	15.35(0.0712)	-9.16(0.0533)	-30.92(0.592)	9.89(1.50)	11.35(3.37)
22	15.52(0.0707)	-9.25(0.0523)	-30.65(0.576)	9.84(1.49)	11.50(3.13)
23	15.69(0.0709)	-9.34(0.0518)	-30.39(0.563)	9.81(1.49)	11.62(2.93)

Zu dieser geometrischen Beschreibung wurden, ebenfalls in der signalnahen Schicht, begriffliche Primitiven abstrahiert, welche versprachlichte Aussagen über Geschwindigkeit, Richtung und andere Informationen über das Fahrverhalten des Fahrzeugs enthalten. Ein sehr kleiner Ausschnitt dieser Datenmenge ist in der bekannten Notation wie folgt angegeben.

1		-i : +i !	agens(obj_10).
1		5 : 95 !	traj_active.
1		5 : 95 !	traj_active(obj_10).
1		-i : +i !	et(obj_10).
1		5 : 95 !	v_Betrag(obj_10, glkl).
1		5 : 95 !	v_Betrag(obj_10, normal).
1		7 : 93 !	modus(obj_10, vorwaerts).
1		5 : 63 !	auf(obj_10, fobj_6).
1		5 : 63 !	f_richtung(obj_10, fobj_6, laengs).
1		64 : 65 !	auf(obj_10, fobj_8).
1		64 : 65 !	f_richtung(obj_10, fobj_8, laengs).
1		25 : 54 !	auf(obj_10, fobj_14).
1		25 : 54 !	f_richtung(obj_10, fobj_14, quer).
1		72 : 95 !	auf(obj_10, fobj_20).
1		72 : 95 !	f_richtung(obj_10, fobj_20, laengs).
1		66 : 71 !	auf(obj_10, fobj_21).
1		66 : 71 !	f_richtung(obj_10, fobj_21, laengs).
1		11 : 65 !	auf(obj_10, fobj_11).
0.52		11 : 11 !	f_richtung(obj_10, fobj_11, laengs).
0.53		12 : 12 !	f_richtung(obj_10, fobj_11, laengs).
0.55		13 : 13 !	f_richtung(obj_10, fobj_11, laengs).
0.57		14 : 14 !	f_richtung(obj_10, fobj_11, laengs).
0.58		15 : 15 !	f_richtung(obj_10, fobj_11, laengs).
0.6		16 : 16 !	f_richtung(obj_10, fobj_11, laengs).
0.61		17 : 17 !	f_richtung(obj_10, fobj_11, laengs).
0.62		18 : 18 !	f_richtung(obj_10, fobj_11, laengs).
0.63		19 : 19 !	f_richtung(obj_10, fobj_11, laengs).
0.64		20 : 20 !	f_richtung(obj_10, fobj_11, laengs).
0.65		21 : 21 !	f_richtung(obj_10, fobj_11, laengs).
0.67		22 : 22 !	f_richtung(obj_10, fobj_11, laengs).
0.68		23 : 23 !	f_richtung(obj_10, fobj_11, laengs).
0.69		24 : 24 !	f_richtung(obj_10, fobj_11, laengs).
0.71		25 : 25 !	f_richtung(obj_10, fobj_11, laengs).



Primitiv-begriffliche Beschreibungen werden nicht nur bezüglich des Fahrverhaltens des Agenten berechnet, sondern auch eine begriffliche Kreuzungsbeschreibung wird auf diese Weise erzeugt. Auch diese Kreuzungsbeschreibung ist unten angegeben. Zum einen werden dort einzelnen Fahrbahnteilen Typbezeichnungen zugeordnet, zum anderen werden räumliche Relationen zwischen Fahrbahnteilen extrahiert.

<p>always rechtspfeil(fobj_30).  always geradeauspfeil(fobj_33).  always linkspfeil(fobj_35).  always linkspfeil(fobj_37).  always fussgaengerfurt(fobj_39).  always haltlinie(fobj_41).  always fussgaengerfurt(fobj_43).  always fussgaengerfurt(fobj_54).  always fussgaengerfurt(fobj_45).  always fussgaengerfurt(fobj_51).  always fussgaengerfurt(fobj_47).  always radweg(fobj_26).  always knotenpunktzentrum(fobj_25).  always schiene(fobj_29).  always fussgaengerfurt(fobj_50).  always fahrspur(fobj_2).  always fahrspur(fobj_4).  always fahrspur(fobj_6).  always fahrspur(fobj_8).  always fahrspur(fobj_13).  always fahrspur(fobj_15).  always fahrspur(fobj_17).  always fahrspur(fobj_19).  always fahrspur(fobj_21).  always fahrspur(fobj_11).  always linksabbiegespur(fobj_5).  always geradeausspur(fobj_3).  always auf(fobj_4, fobj_25).  always vor(fobj_5, fobj_6).  always auf(fobj_6, fobj_25).  always vor(fobj_7, fobj_8).  always auf(fobj_8, fobj_25).</p>	<p>always geradeauspfeil(fobj_32).  always linkspfeil(fobj_34).  always linkspfeil(fobj_36).  always haltlinie(fobj_38).  always fussgaengerfurt(fobj_40).  always fussgaengerfurt(fobj_42).  always fussgaengerfurt(fobj_53).  always haltlinie(fobj_44).  always fussgaengerfurt(fobj_46).  always fussgaengerfurt(fobj_52).  always fussgaengerfurt(fobj_48).  always radweg(fobj_27).  always schiene(fobj_28).  always fussgaengerfurt(fobj_49).  always fahrspur(fobj_1).  always fahrspur(fobj_3).  always fahrspur(fobj_5).  always fahrspur(fobj_7).  always fahrspur(fobj_9).  always fahrspur(fobj_14).  always fahrspur(fobj_16).  always fahrspur(fobj_18).  always fahrspur(fobj_20).  always fahrspur(fobj_10).  always fahrspur(fobj_12).  always linksabbiegespur(fobj_7).  always rechtsabbiegespur(fobj_1).  always aufbringung(fobj_5, fobj_34).  always vor(fobj_6, fobj_20).  always aufbringung(fobj_7, fobj_36).  always vor(fobj_8, fobj_21).  always auf(fobj_38, fobj_3).</p>
---	--

Basierend auf diesen Datenmengen wird die Situationsanalyse durchgeführt. Im Beispiel wurde ein Bericht über das Abbiegeverhalten des betreffenden Fahrzeuges extrahiert. Es wird für jeden Zeitpunkt die aktuelle Situation des Agenten bestimmt.

---

5 : 65 ! turning\_left\_into(obj\_10, ettlinger\_strasse).  
 66 : 95 ! leaving\_intersection(obj\_10).

---

Diese Situationen werden unter Berücksichtigung des gewählten Standpunktes perspektivisch beschrieben.

---

5 : 65 ! turning\_left\_into(obj\_10, ettlinger\_strasse, 9).  
 66 : 95 ! leave(obj\_10, the\_intersection, 10).

---

Dabei stimmen in diesem Fall lediglich aufgrund der gewählten einfachen Darstellung und der Kürze des Beispiels beide Darstellungen ziemlich überein. Mittels der verwendeten Konstruktionsregeln wird diese begriffliche Beschreibung in folgende DRS überführt:

$x_9$ $y_9$ $a_9$ $x_{10}$ $a_{10}$ $\text{obj\_10}(x_9)$ $\text{ettlinger\_strasse}(y_9)$ $a_9$ : <span style="border: 1px solid black; padding: 2px;"><math>x_9</math> turn_left_into <math>y_9</math></span> $x_{10} = x_9$ $\text{the\_intersection}(y_{10})$ $a_{10}$ : <span style="border: 1px solid black; padding: 2px;"><math>x_{10}</math> leave <math>y_{10}</math></span> $a_9 < a_{10}$
--

Aus dieser DRS wird der folgende natürlichsprachliche Text abgeleitet:

**Obj 10 turned left into ettlinger strasse. Then it left the intersection.**

### 5.3 Zusammenfassung

In dieser Arbeit wurde ein Grundstein gelegt zur Ableitung natürlichsprachlicher Beschreibungen des Verhaltens von Fahrzeugen in Realwelt-Straßenverkehrsszenen. Die in der Arbeit hierzu entwickelten Systemkomponenten setzen auf ein bestehendes, am selben Institut entwickeltes Bildfolgenauswertungssystem auf, das entsprechende Videobildfolgen auswertet und daraus geometrische und primitiv-begriffliche Szenenbeschreibungen berechnet. Es wurde festgestellt, daß natürlichsprachliche Beschreibungen von

Abläufen stets perspektivisch sind. Statt nun, wie in anderen vergleichbaren Arbeiten (z.B. [Dance et al. 96], [André et al. 88]), eine oder mehrere Perspektiven der Beschreibung systemintern implizit festzulegen, wurde in der vorliegenden Arbeit angestrebt, dem Benutzer die Wahl der gewünschten Perspektive zu ermöglichen. Dabei wurde eine generische Beschreibung von Erzählerstandpunkten aufgebaut. Damit einher geht die Festlegung dreier Erzählstile (Bericht, szenische Beschreibung aus Protagonisten- bzw. Beobachtersicht) zum Aufbau natürlichsprachlicher Beschreibungen. Beim Entwurf der Systemkomponenten wurde zwischen der Ermittlung des Inhalts einer Beschreibung und dessen Präsentation unterschieden. Erstgenannter Arbeitsschritt wird von der begrifflichen Systemschicht übernommen. Anhand eines bestehenden Formalismus zur Situationsanalyse wird die aktuelle Situation der Hauptakteure ermittelt. Abhängig vom gewählten Standpunkt wird im nächsten Schritt die ermittelte Situationsfolge perspektivisch begrifflich beschrieben. Das Wissen, wie bestimmte Situationen und Situationsfolgen vom gewählten Standpunkt aus zu beschreiben sind, entnimmt das System aus Perspektivierungslisten, die entsprechende Einträge für jeden zu berücksichtigenden Standpunkt enthalten. Ergebnis der Berechnungen der begrifflichen Schicht ist eine perspektivische, begriffliche Beschreibung, die in der Notation unscharf metrisch temporaler Logikfakten den Inhalt der zu generierenden Beschreibung enthält. Es gilt nun, diesen Inhalt in eine natürlichsprachliche Darstellung zu übertragen. Dieser Schritt wird von der natürlichsprachlichen Systemschicht vorgenommen, bei deren Entwicklung hauptsächlich linguistische Fragestellungen zum Tragen kommen (Diskursrepräsentationstheorie (DRT), [Kamp&Reyle 93]). Die DRT, die sich mit der semantischen Repräsentation natürlichsprachlicher Texte beschäftigt, wurde für den in dieser Arbeit verfolgten Ansatz adaptiert. Die Aufgabe der natürlichsprachlichen Schicht besteht hauptsächlich darin, den Stil des Textes zu bestimmen. Dies bedeutet, dafür Sorge zu tragen, daß der in der begrifflichen Schicht ermittelte Inhalt in seiner für den Benutzer verständlichsten Weise natürlichsprachlich formuliert wird. Diesen Anspruch kann die natürlichsprachliche Schicht in ihrer jetzigen Form noch nicht voll erfüllen. Insbesondere ist die DRT noch nicht so weit fortgeschritten, daß ein wohlfundiertes Regelwerk zur Ableitung von Texten aus der rechnerinternen Repräsentation bestünde. Behelfsweise wurden einige wenige solcher Regeln im Rahmen dieser Arbeit entwickelt, die die Ableitung einfacher Sätze ermöglichen. Auf den bestehenden Formalismus der DRT wurde deshalb zugegriffen, weil die bei der Konstruktion einer natürlichsprachlichen Schicht relevant werdenden linguistischen Fragestellungen den Arbeitsbereich klassischer Bildfolgenauswertung überschreiten und sich deshalb die Mitwirkung von Spezialisten in diesem neuen Bereich empfiehlt. Die in dieser Arbeit entwickelten Methoden wurden an zahlreichen Bildfolgen getestet. Einige Ergebnisse sind in Anhang A aufgeführt.

## 5.4 Bewertung

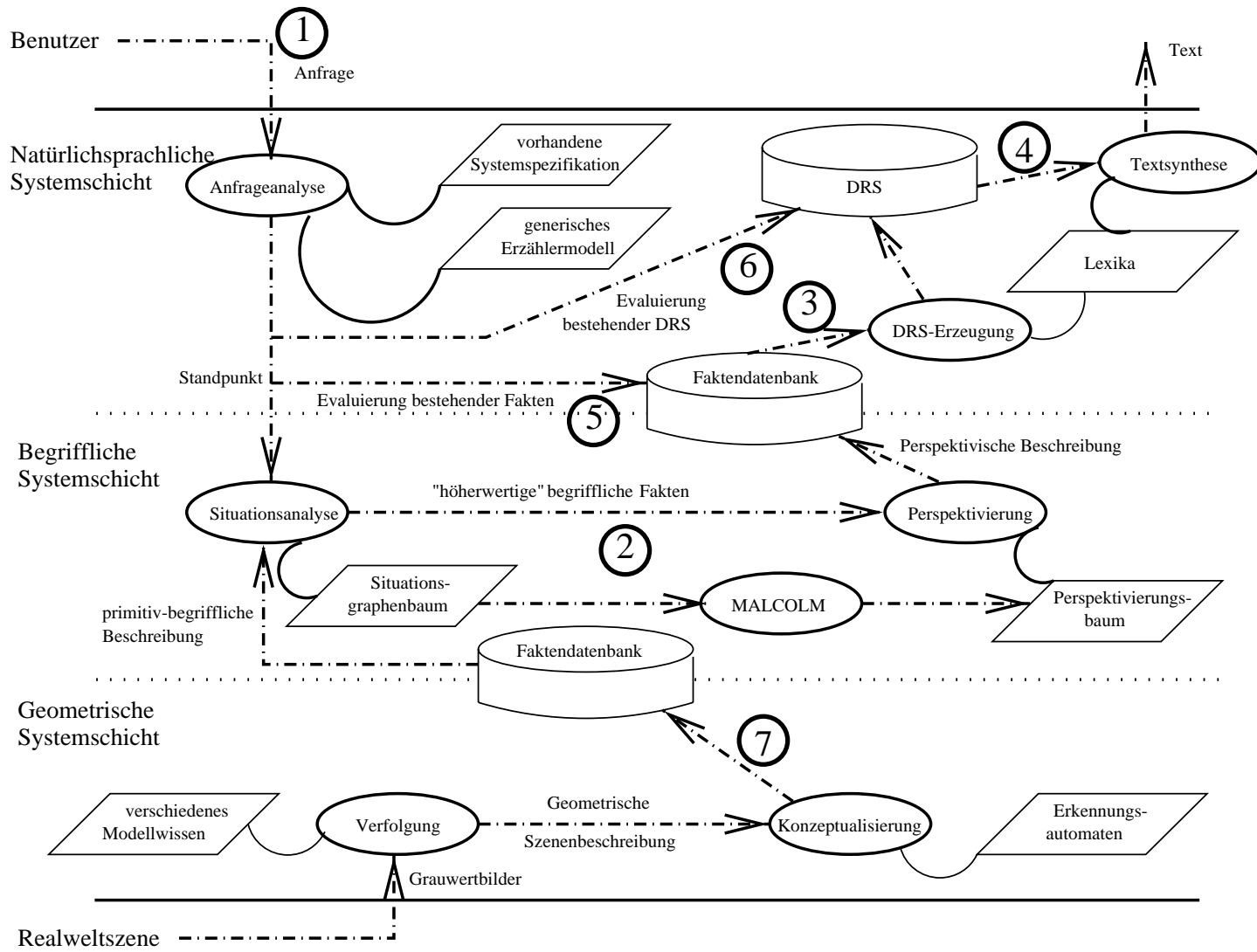
Bei der Berechnung natürlichsprachlicher Beschreibungen kann es zu drei klassischen Fehlerfällen kommen.

- **Fehler in der Verfolgung:** Die erste Klasse bilden Fehler in der signalnahen Auswertung, die sich als Meßfehler äußern. Die vorliegende Arbeit liefert keine Ansätze zur Verbesserung der signalnahen Auswertung. Die Quote der erfolgreichen Fahrzeugverfolgungen liegt derzeit bei etwas über 80 Prozent. Allerdings kann die Transformation der Auswertungsergebnisse in eine begriffliche oder natürlichsprachliche Beschreibung dazu beitragen, daß versteckte Fehler offensichtlicher zu Tage treten. Eine falsche Zahl in einer geometrischen Beschreibung ist schwerer zu entdecken als ein falscher Satz in einer natürlichsprachlichen Beschreibung. Die Analyse, wie der falsche Satz zustande kam, kann dann aber dazu führen, daß diese falsche Zahl, die ansonsten unentdeckt geblieben wäre, entdeckt wird.
- **Fehler in der Modellierung:** Modellierungsfehler treten am häufigsten in der begrifflichen Auswertung auf. Sie werden meistens dann sichtbar, wenn ein Fahrzeug sich nicht so verhält, wie es systemintern erwartet wird. Der gewählte Situationsanalyseformalismus bietet anhand seiner hierarchischen Struktur und seiner begrifflichen Beschreibbarkeit durch unscharf metrisch temporale Logikformeln eine im Vergleich zu anderen Szeneninterpretationsformalismen (z.B. Bayes-Netze in [Intille 98], [Buxton&Gong 95]) bequeme Erweiterbarkeit. Modellierungsfehler können jedoch auch zukünftig nicht ausgeschlossen werden, da die Vielfalt möglichen abweichenden Fahrzeugverhaltens nicht vollständig kalkulierbar ist. Die Situationsanalyse bietet jedoch zumindest die Möglichkeit, abweichendes Verhalten zu erkennen, wenn auch nicht zu beschreiben.
- **Fehler in der natürlichsprachlichen Schicht:** Die Transformation einer begrifflichen Beschreibung in eine DRS liefert keine neue Fehlerquelle. Konstruktionsregeln, Wortschatz und Grammatik wurden so gewählt, daß der Inhalt begrifflicher Beschreibungen frei von Mehrdeutigkeiten in die DRS eingeht. Auch wurden die Regeln für die Ableitung von Text aus der DRS so einfach gehalten, daß, basierend auf dem verwendeten Wortschatz, die Verfälschung des Inhaltes bei der Transformation von DRS in Text stark eingeschränkt werden kann. In den durchgeführten Experimenten wurden bei diesem Auswertungsschritt keine Fehler beobachtet. Unzulänglichkeiten der natürlichsprachlichen Schicht äußern sich demnach ausschließlich in 'schlechten' Formulierungen des Textes. Dieser Mangel wird noch absehbare Zeit Bestand haben, da die Entwicklung der linguistischen Plattform erst weiter voranschreiten muß.

## 5.5 Ausblick

Abbildung 5.2 zeigt noch einmal den Systementwurf, wie er bei der Erstellung dieser Arbeit mittelfristig angestrebt wurde. Selbige Abbildung ist auch im Abschnitt 2.6 angegeben und erläutert. Im folgenden werden die nach Ansicht des Autors unmittelbar

Abbildung 5.2: Angestrebter Systementwurf



anstehenden Weiterentwicklungen seiner Arbeit angeregt. Dazu wird das in der Abbildung dokumentierte Szenario mit dem in dieser Arbeit Erreichten verglichen. Die in der folgenden Aufzählung aufgeführten Nummern beziehen sich auf die entsprechenden Nummern in der Abbildung und sollen dem Leser die Einordnung der jeweils besprochenen Teilkomponente in das Gesamtsystem erleichtern.

1. Angestrebt wird langfristig, daß der Benutzer eine natürlichsprachliche Anfrage stellt, die vom System analysiert wird. Aus dieser soll dann gegebenenfalls im Dialog mit dem Benutzer der zweckmäßige Erzählerstandpunkt ermittelt werden, bezüglich dessen die Antwort zu generieren ist. Im gegenwärtigen Entwurfsstadium spezifiziert der Benutzer seinen Standpunkt noch nicht durch eine natürlichsprachliche Anfrage. Es existiert allerdings ein generisches Erzählermodell, das durch Setzen der entsprechenden freien Parameter ausgeprägt wird. Es wird vom Benutzer bis auf weiteres verlangt, daß er die von ihm gewünschten Parameterwerte explizit angibt (Erzählstil, Hauptakteure, zu berücksichtigende Nebenakteure, besonders zu berücksichtigende Abläufe).
2. Die Berechnung einer 'höherwertigen begrifflichen Beschreibung' (in der Abbildung durch die Ziffer '2' markiert) wurde in der angestrebten Weise realisiert. Der Parametersatz für die Graphtraversierung wird, wie oben angegeben, nicht aus einer natürlichsprachlichen Anfrage abgeleitet, sondern vom Benutzer direkt vorgegeben. Er bildet die Grundlage für die perspektivische Situationsbeschreibung, welche durch Ausprägung des von MALCOLM erzeugten Perspektivierungsbaums erfolgt.
3. Die dabei gewonnenen Fakten werden, wie in der Abbildung gezeigt, in die semantische Repräsentationsform 'DRS' überführt.
4. Daraus wird der Antwort-Text, also die natürlichsprachliche Szenenbeschreibung vom vorgegebenen Standpunkt aus, abgeleitet. Zu diesem Zweck wurden einfache Ableitungsregeln kreiert, welche für die bisherigen Anforderungen ausreichen. Allerdings besteht für die Konstruktion solcher Regeln bisher keine ausreichende Theorie. Eine solche gilt es sukzessive zu entwickeln. Dies ist jedoch in erster Linie eine linguistische Aufgabenstellung.
5. Diese Ziffer markiert eine der beiden 'Abkürzungen', die das System bei der Textgenerierung zur Vermeidung redundanter Berechnungen ermöglichen soll. Dabei wird beabsichtigt, die bereits vor der aktuellen Anfrage generierten Fakten dahingehend zu überprüfen, ob sie bereits zur Anfragebeantwortung ausreichen oder sich zumindest in Teilen verwenden lassen. Dieser Zweig wird in der vorliegenden Arbeit nicht verfolgt. Seine Installierung wird vom Autor nicht als dringlichste Erweiterungsaufgabe angesehen, da sein praktischer Nutzen für die Entwurfsphase, in der sich die natürlichsprachliche Auswertung noch befindet, für nicht allzu groß eingeschätzt wird.

6. Diese Ziffer verweist auf die zweite denkbare ‘Abkürzung’. Dabei ist vorgesehen, eine bereits vorhandene DRS dahingehend zu überprüfen, ob die in ihr enthaltene Information zur Antwortgenerierung ausreicht. Zur Gewährleistung, daß nur pragmatische DRSen erzeugt werden (siehe Abschnitt 4.5.1), wird bisher vor jeder neuen Anfrage eine bereits bestehende DRS gelöscht. Dies geschieht, da bei der etwaigen Vereinigung zweier pragmatischer DRSen nicht zwangsläufig wieder eine pragmatische DRS erzeugt wird. Desweiteren existiert bisher kein Verfahren, wie durch Manipulation einer bestehenden, nicht pragmatischen DRS eine pragmatische erzeugt werden kann (vgl. Abschnitt 4.6). Aus diesen Gründen kann die angestrebte ‘Abkürzung’ derzeit nicht realisiert werden. Ihre Schaffung wird vom Autor wiederum nicht als ein zentrales Thema kurzfristig anzustrebender Erweiterungen gesehen.
  
7. Die Kommunikation zwischen geometrischer Schicht und den beiden aufsetzenden Schichten erfolgt bisher nur in einer Richtung und durch Austausch von Dateien (siehe [Haag&Nagel 98b] für eine Ausnahme von dieser Regel, wenn begriffliches Wissen ausgenutzt wird, um Verhaltensänderungen vollständig verdeckter Fahrzeuge bei ihrer Verfolgung zu erkennen). Mit anderen Worten bedeutet dies, daß die beiden neu geschaffenen Systemschichten noch nicht in das Gesamtsystem integriert sind. Diese Integration ist die dringlichste Aufgabe, die sich nach Abschluß der vorliegenden Arbeit stellt. Diese Integration ermöglicht, daß die Auswertung von Bilddaten durch alle Auswertungsstufen streng inkrementell erfolgen könnte. Bisher erfolgt die Auswertung im Rahmen dieser Arbeit nicht streng inkrementell: Zuerst wird die geometrische Szenenbeschreibung komplett über alle Zeitpunkte berechnet. Erst dann erfolgt die Situationsanalyse (wiederum für alle Zeitpunkte). Erst nach deren vollständigem Abschluß erfolgt die komplette DRS-Berechnung, dann wiederum erst die Texterzeugung.

Eine weitere, sehr wesentliche Erweiterungsmöglichkeit wird aus der Abbildung nicht unmittelbar deutlich. Aufgabe der natürlichsprachlichen Schicht soll es in erster Linie sein, eine stilistisch ansprechende Versprachlichung des zuvor ermittelten Inhalts zu ermöglichen. Mit anderen Worten unterliegt die Textplanung der Verantwortung der natürlichsprachlichen Schicht. Dies ist zum gegenwärtigen Stand noch nicht vollständig erreicht. Ein Beispiel soll dies verdeutlichen.

Für ein Fahrzeug werden nacheinander folgende Situationen erkannt:

1	Wechseln_auf_Ausweichspur
2	Gleichziehen_mit_anderem_Fahrzeug
3	Einscheren_vor_anderem_Fahrzeug
4	Verlassen_Zapfsäule

Vom gewählten Standpunkt aus wird aus Situation\_1 die Absicht abgeleitet, daß das Agens überholen will. In Situation\_2 wird erkannt, daß das Agens aktuell überholt.

Situation\_3 sagt aus, daß der Überholvorgang abgeschlossen wurde. Dadurch läßt sich die Situationsfolge 1-4 wie folgt natürlichsprachlich beschreiben (in englischer Sprache, das Agens werde als 'Object\_4' bezeichnet):

*Object\_4 changed to the passing lane in order to overtake another car. Then it overtook the car. Later, overtaking was completed. Then it left the filling station.*

Zum Vergleich betrachte man folgende natürlichsprachliche Beschreibung:

*Object\_4 changed to the passing lane in order to overtake another car. Then it left the filling station.*

Dieser Text beschreibt den gleichen Sachverhalt wie der erstgenannte. Er ist kürzer, stilistisch besser und vermittelt im Vergleich zu erstgenanntem die gleiche Information. Die Formulierung der bloßen Absicht, überholen zu wollen, reicht im Beschreibungskontext der zweiten Beschreibung aus, um beim Leser den gewünschten Eindruck zu erwecken, daß diese Absicht schließlich auch umgesetzt wurde. Eine explizite Erwähnung der erfolgten Umsetzung wie in erstgenannter Beschreibung ist folglich redundant. Da die Redundanzvermeidung in einer Beschreibung als stilistische Aufgabe betrachtet wird, ist für ihre Erkennung die natürlichsprachliche Schicht verantwortlich. Die bisher vorhandenen Methoden reichen jedoch nicht aus, Redundanz in einer DRS zu eliminieren. Sie reichen auch nicht aus, um diese Redundanz bei der Ableitung eines Textes aus der DRS zu eliminieren. Eine Realisierung zumindest einer der beiden angedeuteten Erweiterungen ist langfristig anzustreben. Auch diese Aufgabenstellung greift in linguistische Forschungsbereiche und sollte deshalb in Kooperation mit solchen Forschungsgruppen erfolgen.

Bisher wird oben beschriebene Art der Redundanz dadurch vermieden, daß die Textplanung nicht dem System (d.h. der natürlichsprachlichen Schicht) unterliegt, sondern noch im wesentlichen dem Systementwerfer bei der Implementierung der Perspektivierungstabellen (vgl. Abschnitt 3.4.4). Diese Tabellen enthalten das Wissen, wie bestimmte Situationen aus bestimmten Perspektiven zu beschreiben sind. Die Tabellengestaltung ist vom Entwerfer so 'geschickt' vorzunehmen, daß besagte Redundanzen nicht auftreten können. Im Hinblick auf das oben genannte Beispiel muß er berücksichtigen, daß die Durchführung einer Aktion nicht versprachlicht werden muß, wenn unmittelbar zuvor die Absicht versprachlicht wurde, daß besagte Aktion durchgeführt werden soll.

Fazit: Mittelfristig ist dafür zu sorgen, daß der Entwerfer sukzessive aus seiner Verantwortung für die Textplanung entlassen und diese vermehrt auf das System übertragen wird. Diese Zielsetzung erfordert noch hohen Forschungsaufwand.

Die Textplanung bedingt nicht nur Redundanzvermeidung. Sie muß weitere stilistische Entscheidungen ermöglichen können, die an dieser Stelle nicht abschließend aufgezählt werden. Als ein Beispiel sei die Wahl der Zeitform für das Verb eines Satzes genannt. Grundsätzlich wird bisher in jedem Satz die Vergangenheitsform ('Simple Past') verwendet. Die einzelnen Erzählstile erlauben (oder erfordern sogar), daß auch andere Zeitformen verwendet werden (Verlaufsformen, Perfekt, Versprachlichung von Absich-



ten nicht mit ‘want\_to’, sondern durch Verwendung des Futurs etc.). Der Grund, warum andere Zeitformen als die gewählte Vergangenheitsform bisher nicht verwendet werden, liegt nicht darin, daß ihre Bildung zu schwierig wäre. Vielmehr liegt der Grund darin, daß die Möglichkeit, andere Zeitformen zu verwenden, die Entscheidungskompetenz voraussetzt, diese auch angemessen einzusetzen. Auch hierfür ist eine systemgesteuerte Textplanung zu entwickeln.

## 5.6 Andere Diskursbereiche

Es sei festgestellt, daß vorliegende Arbeit einen nützlichen Beitrag zur Berechnung perspektivischer, natürlichsprachlicher Beschreibungen von Straßenverkehrsszenen durch Bildfolgenauswertung leistet. Offen ist bisher die Frage, wie sich die Methoden der vorliegenden Arbeit auf andere Anwendungsgebiete übertragen lassen. Hierzu seien noch einmal die Eckpunkte des am Institut für Algorithmen und Kognitive Systeme entwickelten Bildfolgenauswertungssystems aufgeführt:

1. Robuste signalnahe Bildfolgenauswertung und Berechnung einer hinreichend umfassenden geometrischen Szenenbeschreibung (GSB);
2. Assoziation begrifflicher Primitiven zur GSB;
3. Modellierung im Diskursbereich üblicherweise auftretender Situationen mit Hilfe hierarchischer Situationsgraphenbäume;
4. Abstraktion einer begrifflichen Beschreibung durch Assoziation der begrifflichen Beschreibung zu dem Schema, das durch den Situationsgraphenbaum gegeben ist, mittels unscharf metrisch temporaler Logik;
5. Automatische Berechnung von Perspektivierungsgraphen in Abhängigkeit zu einem gegebenen Standpunkt zur Ableitung perspektivischer Beschreibungen;
6. Berechnung natürlichsprachlicher Beschreibungen durch DRT aus der perspektivischen, begrifflichen Beschreibung.

Es wird behauptet, daß die in der vorliegenden Arbeit entwickelten Methoden auch auf andere Diskursbereiche, zu denen natürlichsprachliche Szenenbeschreibungen berechnet werden sollen, übertragbar sind. Einer Änderung des Diskursbereichs soll idealerweise hauptsächlich mit der Änderung des systeminternen Modellwissens begegnet werden, beispielsweise bei der Verfolgung sich bewegender Objekte durch die Ersetzung von Fahrbahnmodellen mit Modellen für diejenigen Objekte (z.B. Flugzeuge), die in dem alternativen Diskursbereich behandelt werden. Selbiges gilt für den Bereich der Abstraktion. Hierbei sollen Diskursbereichsänderungen in erster Linie mit einer Änderung des Situationsgraphenbaums und der zugrundeliegenden Terminologie begegnet

werden. Zur Berechnung perspektivischer Beschreibungen für einen anderen Diskursbereich ist dementsprechend eine Anpassung des Perspektivierungsmodells erforderlich. Man betrachte als Beispiel den Bereich der automatischen Videodatenverwaltung. Hierbei steht man vor dem Problem, für den Menschen schwer überschaubare Mengen von Videofilmen zu archivieren und gezielt wiederzufinden. Eine Methode könnte darin bestehen, durch Bildfolgenauswertung natürlichsprachliche Kurzbeschreibungen des Inhalts der Videofilme zu erstellen. Solche Kurzbeschreibungen unterlägen hinsichtlich Perspektivität und Ausführlichkeit vergleichbaren Anforderungen wie bei den in der vorliegenden Arbeit vorgenommenen Untersuchungen. Die Implementierung hierarchischer Situationsgraphenbäume begünstigt insbesondere die Berechnung unterschiedlich ausführlicher Beschreibungen.

Bereits in der Einleitung zu der vorliegenden Arbeit wurde festgestellt, daß durch die rechnergesteuerte Abarbeitung von Algorithmen generell Daten berechnet werden, deren Versprachlichung für bestimmte Benutzer interessant und hilfreich sein könnte. Dies gilt auch außerhalb der Bildfolgenauswertung. So wird beispielsweise im Rahmen des Mensch-Maschine-Dialogs nach Methoden geforscht, wie dem Benutzer, soweit er daran interessiert ist, automatisch natürlichsprachlich dargestellt werden kann, welche ansonsten 'versteckten' Operationen das System oder Betriebssystem, mit dem er arbeitet, gerade durchführt. Zwar wurden dementsprechende Experimente bisher noch nicht vollzogen, doch erscheint es zumindest möglich, daß selbst für solche Anwendungsgebiete Methoden der vorliegenden Arbeit verwendet werden können. Wenn es gelänge, interne Zustände des beobachteten Systems und deren Zusammenhänge begrifflich zu beschreiben, so böte dies die Grundlage zur Erstellung eines entsprechenden Situationsgraphenbaums. Zur Festlegung von Art und Umfang der Ableitung natürlichsprachlicher, perspektivischer Beschreibungen würden dann wieder die Überlegungen aus der vorliegenden Arbeit relevant.

Zusammengefaßt wird behauptet, daß die vorliegende Arbeit diskursübergreifend Relevanz bei der Versprachlichung automatisch erzeugter Information besitzt. Dreh- und Angelpunkt bei der Diskursbereichsanpassung ist dabei das dem System zur Verfügung stehende Weltwissen in Form von Situations- und Perspektivierungsgraphen. Die Bereitstellung solcher Modelle bedeutet umfangreiche Handarbeit durch (menschliche) Spezialisten. Es stellt sich abschließend die Frage, wie diese Handarbeit eingeschränkt werden kann. Zum einen kann dabei an die Automatisierung der Modellerstellung gedacht werden. Die Einbindung entsprechender Lernverfahren kann hierbei möglicherweise hilfreiche Unterstützung bieten. Entsprechende Untersuchungen wurden bisher jedoch noch nicht vorgenommen. Mindestens ebenso aussichtsreich erscheint jedoch der Ansatz, der in der bereits weiter oben zitierten Arbeit [Nagel 99] vorgestellt wurde. Dieser Autor regt an, DRSen als Schemata für die Ausprägung zu verwenden. Im Resultat bedeutet dies, daß das Modellwissen, das bisher in Situationsgraphenbäumen enthalten ist, zukünftig natürlichsprachlich als Text formuliert werden soll. Mit Hilfe der DRT wird aus diesem Text automatisch ein Schema erzeugt, anhand dessen die Ausprägung vorgenommen werden kann. Auf diese Weise bleibt die Modellerstellung zwar weiterhin

weitestgehend Handarbeit. Statt eines Spezialisten, der einen entsprechenden Graphen erstellt, wird jetzt aber nur ein natürlichsprachlicher Text benötigt, der entsprechende Informationen enthält. Dieser kann auch von Menschen erstellt werden, die nicht in den internen Aufbau des Systems eingeweiht sind. Die Implementierung des Modellwissens in das System wird dadurch erheblich vereinfacht.

# Anhang A

## Auswertungsergebnisse

Im Folgenden sind einige aktuelle Auswertungsergebnisse aufgeführt, die auf der Basis der in dieser Arbeit entwickelten und beschriebenen Methoden berechnet wurden. Einzelheiten über Art und Weise der Durchführung dieser Experimente entnehmen man den Erläuterungen aus Abschnitt 5.1. Die ersten Ergebnisse beziehen sich auf Aufnahmen an einer Tankstelle (*Tankstellenszene*). Für das Verhalten jedes verfolgten Fahrzeugs wurde dabei eine szenische Beschreibung aus Protagonisten- und Beobachtersicht sowie ein Bericht generiert. Die jeweiligen Beschreibungen stellen nur das Verhalten eines Fahrzeugs in den Vordergrund. Das Verhalten anderer Fahrzeuge wird nur dann erwähnt, wenn es das Verhalten des Hauptakteurs beeinflusst.

### A.1 Tankstellenszene

#### A.1.1 Objekt 2

**Manuell erstellte Beschreibung:** Objekt 2 fährt von rechts auf der unteren Tankspur, hält an der mittleren Tanksäule und verläßt schließlich die Tankstelle über die linke Ausfahrt (Abbildung A.2).

**Bericht:**

Obj 2 got petrol. It entered the filling station. All filling places were free. Obj 2 chose the lower filling lane. Then it stopped on the second filling place. It took in petrol. It left the filling station.
--

Beurteilung: Der Sachverhalt wurde korrekt erfaßt. Die wesentliche Textaussage, nämlich daß das Fahrzeug tankte, wurde an den Beginn des Textes gestellt. Dies ist möglich, da es sich hier um den Beschreibungsstil ‘Bericht’ handelt, der den beobachteten Ab-



Abbildung A.1: Objekt\_2 der Tankstellenszene.

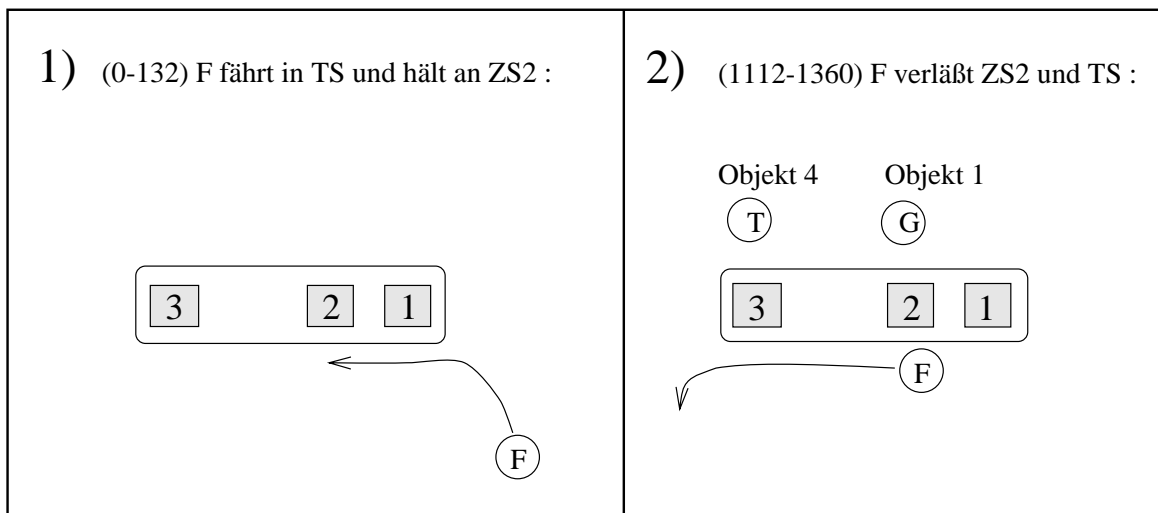


Abbildung A.2: Fahrmanöver von Objekt\_2 (entnommen aus [Haag&Nagel 97]).

lauf im Nachhinein bewertet und beschreibt. Danach wird der Ablauf chronologisch geschildert. Die Formulierung des Textes weist stilistische Mängel auf, die zum jetzigen Stand der Entwicklung unvermeidbar sind. Allerdings wird die chronologische Schilderung des Ablaufs durch die gewählten Formulierungen deutlich und auch an einer Stelle explizit durch die korrekt plazierte Einfügung des Temporaladverbs ‘then’ betont. Im zweiten Satz wurde statt der expliziten Benennung des Agens als Subjekt ein Pronomen verwendet, dessen Referenzierung sich aus dem Text korrekt und eindeutig ergibt. Der dritte Satz gibt korrekt den aktuellen Belegungszustand der Zapfsäulen in Form einer quantorisierten Aussage wieder. Die Angabe des Belegungszustandes erfolgt an dieser Stelle, weil er für die Entscheidung des Fahrers, welche Zapfsäule er ansteuert, wesentlich ist. Gleichzeitig wird durch diese Angabe dem Empfänger der Beschreibung, der die Szene selbst möglicherweise nicht sieht, ein Bild von der Verkehrssituation vermittelt, wie sie sich dem Agenten darstellt.

### **Szenische Beschreibung aus Protagonistensicht:**

I needed petrol. I, therefore, entered the filling station. All filling places were free. I chose the lower filling lane. Then I stopped on the second filling place. I took in petrol. I left the filling station.
---

Beurteilung: Hier wurde dieselbe Szene aus einer anderen Perspektive, nämlich vom Standpunkt des Agenten, beschrieben. Die Beschreibung ist inhaltlich korrekt. Die Formulierung des ersten und zweiten Satzes wird aus der Beobachtung, daß das Fahrzeug auf die Tankstelle fährt, hergeleitet. Im wesentlichen stimmt die Formulierung des Textes mit der Formulierung des Berichts überein.

### **Szenische Beschreibung aus Beobachtersicht:**

Obj 2 entered the filling station. Then it drove on the lower filling lane. Then it stopped on the second filling place. It took in petrol. It left the filling station.
--

Beurteilung: Auch diese Beschreibung ist inhaltlich korrekt. Da der Beobachter den Ablauf parallel zu dessen Verlauf beschreibt und nur wenig Wissen über die innere Handlungsmotivation des Agenten besitzt, enthält seine Schilderung nur wenige über die reine Beobachtung hinausgehende Interpretationen. Es wird festgestellt, daß das Fahrzeug auf die Tankstelle fährt. Über die Hintergründe dieser Aktion wird, im Gegensatz zum Bericht und der Beschreibung aus Protagonistensicht, nichts ausgesagt, da sie größtenteils spekulativ wären. So erfolgt auch keine Aussage über den aktuellen Belegungszustand der Zapfsäulen beim Eintreffen des Fahrzeugs. Der Beobachter beschreibt lediglich die Veränderungen in der Szene, die von ihm wahrgenommen werden. Aus dem gleichen Grund wird statt der Formulierung ‘Agens wählt die untere Tankspur’

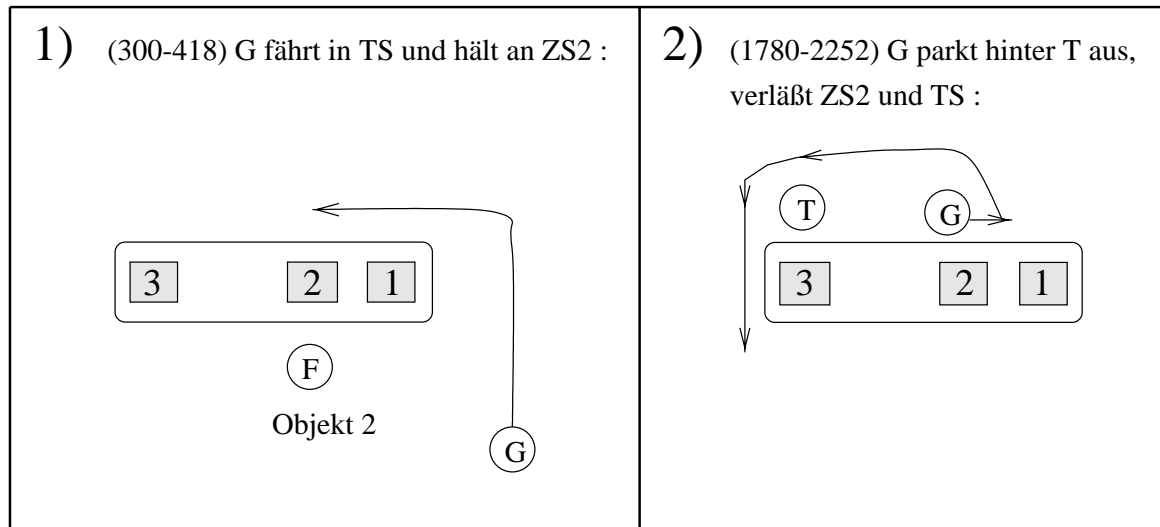


Abbildung A.3: Fahrmanöver von Objekt\_1 (entnommen aus [Haag&Nagel 97]).

die neutralere Formulierung ‘Agens fährt auf der unteren Tankspur’ gewählt. Letztgenannte Formulierung ergibt sich aus der reinen Beobachtung eher als erstgenannte, die ein geplantes Handeln des Agens impliziert, über dessen Ausmaß der Beobachter wenig Vorab-Kenntnis besitzt.

### A.1.2 Objekt 1

**Manuell erstellte Beschreibung:** Objekt 1 fährt von der rechten Einfahrt aus hinter die mittlere Tanksäule und steht dort für ca. 1300 Halbbildzeitpunkte, also 26 Sekunden. Dann setzt es etwas zurück, um schließlich wieder vorwärts anzufahren und nach dem Umfahren des zwischenzeitlich hinter der linken Tanksäule stehengebliebenen Objekts 4 die Tankstelle zu verlassen (Abbildung A.3).

**Bericht:**

Obj 1 got petrol. It entered the filling station. On the lower filling lane, one filling place was occupied. Obj 1 chose the upper filling lane. Then it stopped on the second filling place. It took in petrol. It left the dispensing pump. It, therefore, had to overtake another car which stood on the last filling place. First it backed up. Then it passed another car. Then it left the filling station.

Beurteilung: Bei der Bewertung dieses und der folgenden Testergebnisse werden Aussagen, die schon bei der Bewertung vorheriger Auswertungsergebnisse gemacht wurden,



Abbildung A.4: Objekt\_1 der Tankstellenszene.

nicht wiederholt, auch wenn sie für das jeweils neu bewertete Ergebnis ebenfalls zutreffen. Die extrahierte Beschreibung des Verhaltens von 'Objekt\_1' ist inhaltlich korrekt, auch wenn die Beschreibung des Rücksetzmanövers vor dem Überholen des anderen Fahrzeugs stilistisch nicht übermäßig geglückt ist. Durch die korrekte Verwendung des Kausaladverbs *therefore* konnte ein Bezug zwischen verschiedenen Beobachtungen hergestellt werden.

#### **Szenische Beschreibung aus Protagonistensicht:**

I needed petrol. I, therefore, entered the filling station. On the lower filling lane, one filling place was occupied. So, I chose the upper filling lane. Then I stopped on the second filling place. I took in petrol. I finished getting petrol. I, therefore, wanted to leave the filling station. So, I had to overtake another car. So, I backed up. Then I changed to the passing lane. Then I passed another car. Then I left the filling station.



Beurteilung: Wiederum handelt es sich um eine inhaltlich plausible Beschreibung. Die Wahl der Fahrbahn wurde als Konsequenz aus dem aktuellen Belegungszustand interpretiert. Dies wurde durch die einleitende Verwendung des Wortes 'so' zu Beginn des vierten Satzes herausgestellt.

**Szenische Beschreibung aus Beobachtersicht:**

Obj 1 entered the filling station. Then it drove on the upper filling lane. Later obj 1 stopped on the second filling place. It took in petrol. Obj 1 backed up. Then it changed to the passing lane. Then it passed another car. Then it left the filling station.

Beurteilung: Auch hier handelt es sich um eine korrekte Beschreibung.

**A.1.3 Objekt 4**

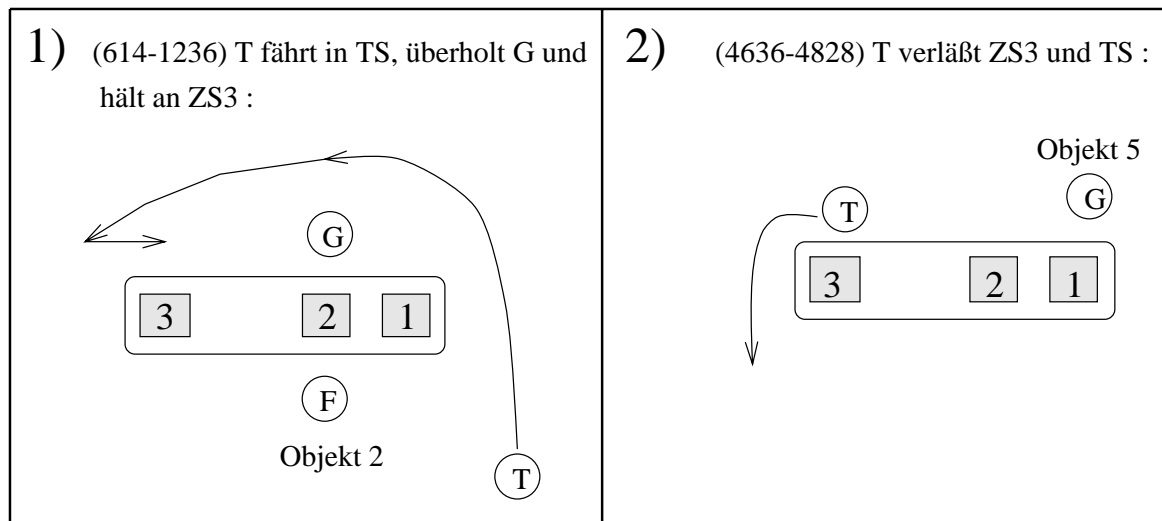


Abbildung A.5: Fahrmanöver von Objekt\_4 (entnommen aus [Haag&Nagel 97]).

**Manuell erstellte Beschreibung:** Objekt 4 weicht von rechts kommend dem hinter der mittleren Tanksäule stehenden Objekt 1 aus, schert vor ihm ein, fährt bis zum linken Bildrand, bleibt dort stehen, setzt zurück und hält schließlich vor dem stehenden Objekt 1 hinter der linken Tanksäule. Nach dem Stehen über 3400 Halbbilder (68 Sekunden) verläßt Objekt 4 die Tankstelle über die linke Ausfahrt (vgl. Abbildung A.5).



Abbildung A.6: Trajektorie von Objekt\_4 der Tankstellenszene (in der Momentaufnahme hinter dem Werbemast befindlich).

#### Bericht:

Obj 4 got petrol. It entered the filling station. On the lower filling lane, one filling place was occupied. On the upper filling lane, one filling place was occupied. Obj 4 chose the upper filling lane. Then it headed for the last filling place. It, therefore, had to overtake another car which stood on the second filling place. Then it changed to the passing lane. Then it passed another car. Then it changed back to the filling lane. Then it stopped behind the last filling place in order to back up to the last filling place. Then it stopped on the last filling place. It took in petrol. It left the filling station.

Beurteilung: Die Beschreibung ist inhaltlich korrekt, in ihrer Formulierung jedoch deutlich verbesserungsfähig. In der Verwendung des zielbeschreibenden Verbes *to\_head\_for* wurde einmal mehr die berichterstatterische Stilistik in der Beschreibung deutlich.

### **Szenische Beschreibung aus Protagonistensicht:**

I needed petrol. I, therefore, entered the filling station. On the lower filling lane, one filling place was occupied. On the upper filling lane, one filling place was occupied, too. I chose the upper filling lane. Then I wanted to overtake another car. I, therefore, changed to the passing lane. Then I passed another car. Then I changed back to the filling lane. Then I stopped behind the last filling place. So, I had to back up to the last filling place. I stopped on the last filling place. I took in petrol. I left the filling station.

### **Szenische Beschreibung aus Beobachtersicht:**

Obj 4 entered the filling station. Then it drove on the upper filling lane. Later obj 4 changed to the passing lane. Then it overtook another car. Then it changed back to the filling lane. Then it stopped behind the last filling place. Then it backed up to the last filling place. Then it stopped on the last filling place. It took in petrol. It left the filling station.

## **A.1.4 Objekt 10**

**Manuell erstellte Beschreibung:** Objekt 10 fährt von rechts ins Blickfeld der aufnehmenden Kamera und bleibt hinter der rechten Tanksäule stehen. Nach einem Aufenthalt von etwa 1200 Halbbildzeitpunkten fährt es wieder an, überholt das hinter der linken Tanksäule stehende Objekt 4 und verläßt die Tankstelle über die linke Ausfahrt (vgl. Abbildung A.8).

### **Bericht:**

Obj 10 got petrol. It entered the filling station. On the upper filling lane, one filling place was occupied. Obj 10 chose the upper filling lane. Then it stopped on the first filling place. It took in petrol. It left the dispensing pump. It, therefore, had to overtake another car which stood on the last filling place. Then it passed another car. Then it left the filling station.



Abbildung A.7: Objekt\_10 der Tankstellenszene.

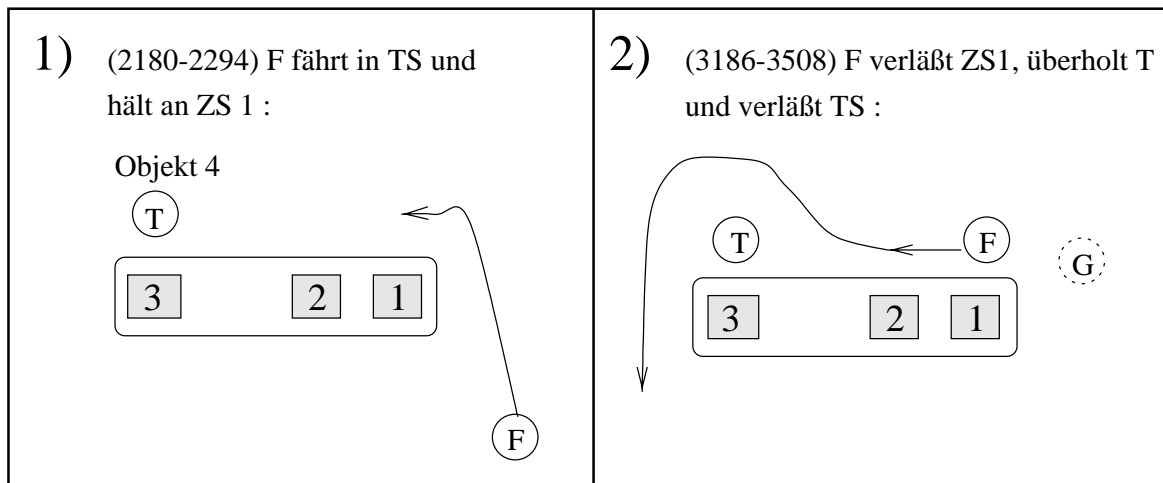


Abbildung A.8: Fahrmanöver von Objekt\_10 (entnommen aus [Haag&Nagel 97]).

### **Szenische Beschreibung aus Protagonistensicht:**

I needed petrol. I, therefore, entered the filling station. On the upper filling lane, one filling place was occupied. I chose the upper filling lane. Then I stopped on the first filling place. I took in petrol. I finished getting petrol. I, therefore, wanted to leave the filling station. So, I had to overtake another car which stood on the last filling place. So, I changed to the passing lane. I passed another car. Then I left the filling station.

### **Szenische Beschreibung aus Beobachtersicht:**

Obj 10 entered the filling station. Then it drove on the upper filling lane. Then it stopped on the first filling place. It took in petrol. Obj 10 changed to the passing lane. Then it passed another car. Then it left the filling station.

## **A.1.5 Objekt 5**

**Manuell erstellte Beschreibung:** Objekt 5 fährt von rechts hinter die rechte Tanksäule und hält dort an. Nach etwas über 2000 Halbbildern setzt es etwas zurück, um dem vor ihm rückwärts einparkenden Objekt 3 Platz zu machen. Nach weiteren 400 Halbbildzeitpunkten setzt Objekt 5 noch weiter zurück, um das inzwischen vor ihm stehende Objekt 3 zu umfahren und die Tankstelle über die linke Ausfahrt zu verlassen (Abbildung A.10).

### **Bericht:**

Obj 5 got petrol. It entered the filling station. On the upper filling lane, one filling place was occupied. Obj 5 chose the upper filling lane. Then it stopped on the first filling place. It took in petrol. It left the filling station.

### **Szenische Beschreibung aus Protagonistensicht:**

I needed petrol. I, therefore, entered the filling station. On the upper filling lane, one filling place was occupied. I chose the upper filling lane. Then I stopped on the first filling place. I took in petrol. I left the filling station.

### **Szenische Beschreibung aus Beobachtersicht:**

Obj 5 entered the filling station. Then it drove on the upper filling lane. Then it stopped on the first filling place. It took in petrol. It left the filling station.



Abbildung A.9: Objekt\_5 der Tankstellenszene.

Beurteilung: Wie die gestrichelte Linie in Abbildung A.10 andeutet, gelang es bisher nicht, das Fahrzeug 'Objekt\_5' über den gesamten Zeitraum stabil zu verfolgen. Beim Wiederanfahren des Fahrzeugs scheitert die Verfolgung bisher noch aufgrund der schwierigen Auswertungsbedingungen (geringer Kontrast, Verdeckungen durch andere Fahrzeuge). Dies bedeutet für die Beschreibung, daß das Bestreben, die Tankstelle wieder zu verlassen, zwar noch erkannt wurde. Allerdings gelang es nicht, den Verlauf des Verlassens der Tankstelle (Zurücksetzen und Überholen eines anderen Fahrzeuges) zu beschreiben. Dieses Beispiel wurde hier berücksichtigt, weil es auch Bestandteil des folgenden Auswertungsergebnisses ist, welches das Verhalten mehrerer Hauptakteure in einem Bericht beschreibt.

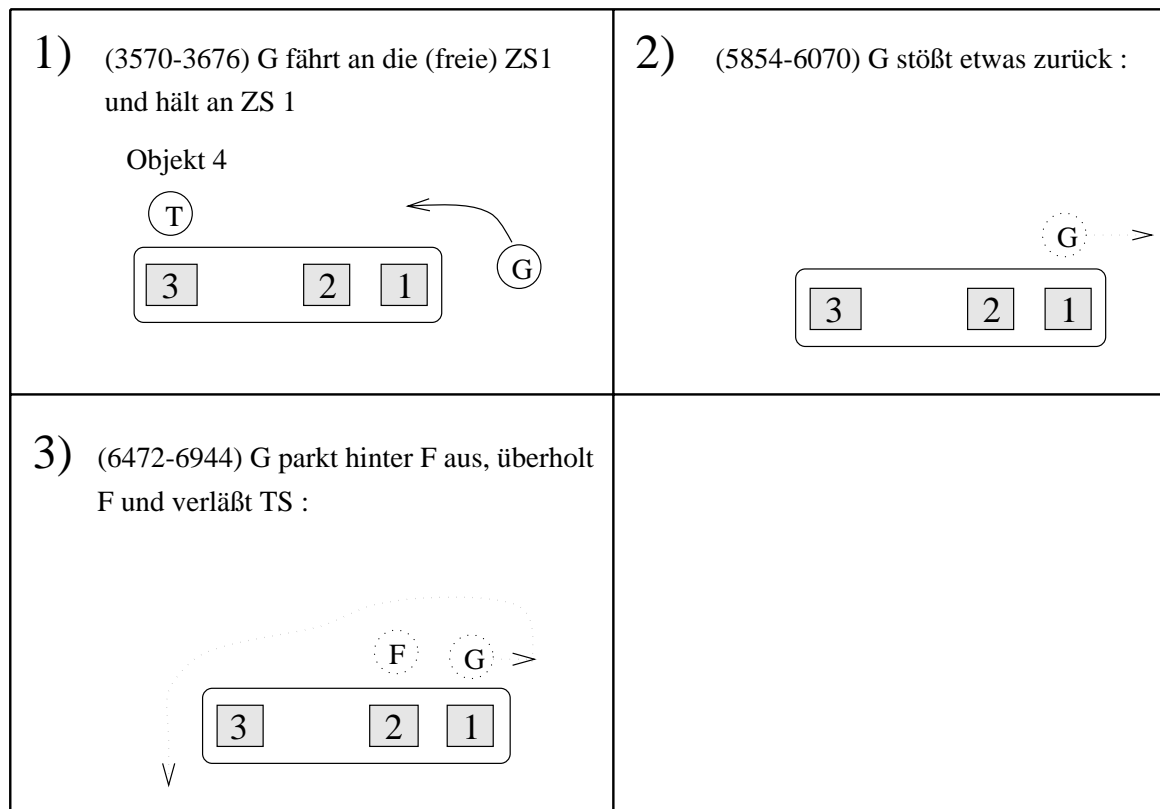


Abbildung A.10: Fahrmanöver von Objekt\_5 (entnommen aus [Haag&Nagel 97]).

### A.1.6 Beschreibung mit mehreren Hauptakteuren

Beim folgenden Auswertungsergebnis wurden das Verhalten der oben bereits behandelten Fahrzeuge mit den Objektnummern 2, 1, 4, 5 in einer Beschreibung zusammengefaßt. Dies geschah, indem für jedes Fahrzeug unabhängig von den anderen eine Situationsanalyse durchgeführt wurde. Daraus wurde jeweils unabhängig voneinander, aber mit derselben Perspektive, eine perspektivische begriffliche Beschreibung abgeleitet. Diese vier Beschreibungen wurden vereint. Daraus ergaben sich, wie das unten aufgeführte Endergebnis zeigen wird, erweiterte Schlußfolgerungsmöglichkeiten. Sie erlaubten, indefinite Bezeichnungen der Bezugsobjekte (*'another car'*) in den oben stehenden Einzelbeschreibungen durch den Eigennamen des jeweils so bezeichneten Fahrzeugs zu ersetzen. Hier der berechnete Bericht:

Obj 2 entered the filling station. All filling places were free. Obj 2 chose the lower filling lane. Then it stopped on the second filling place. It took in petrol. Obj 1 entered the filling station. On the lower filling lane, one filling place was occupied. Obj 1 chose the upper filling lane. Then it stopped on the second filling place. It took in petrol. Obj 4 entered the filling station. On the upper filling lane, one filling place was occupied. Obj 4 chose the upper filling lane. Then obj 4 headed for the last filling place. It had to overtake obj 1. It changed to the passing lane. Then it passed obj 1. Then it changed back to the filling lane. Then it stopped behind the last filling place. Then it backed up to the last filling place. In the meantime obj 2 left the filling station. In the meantime obj 4 stopped on the last filling place. It took in petrol. Obj 1 left the dispensing pump. It had to overtake obj 4. It backed up. Then it passed obj 4. Then it left the filling station. Later obj 5 entered the filling station. On the upper filling lane, one filling place was occupied. Obj 5 chose the upper filling lane. Then it stopped on the first filling place. It took in petrol. Obj 4 left the filling station. Later obj 5 left the filling station.

## A.2 Alte Durlacher-Tor-Szene

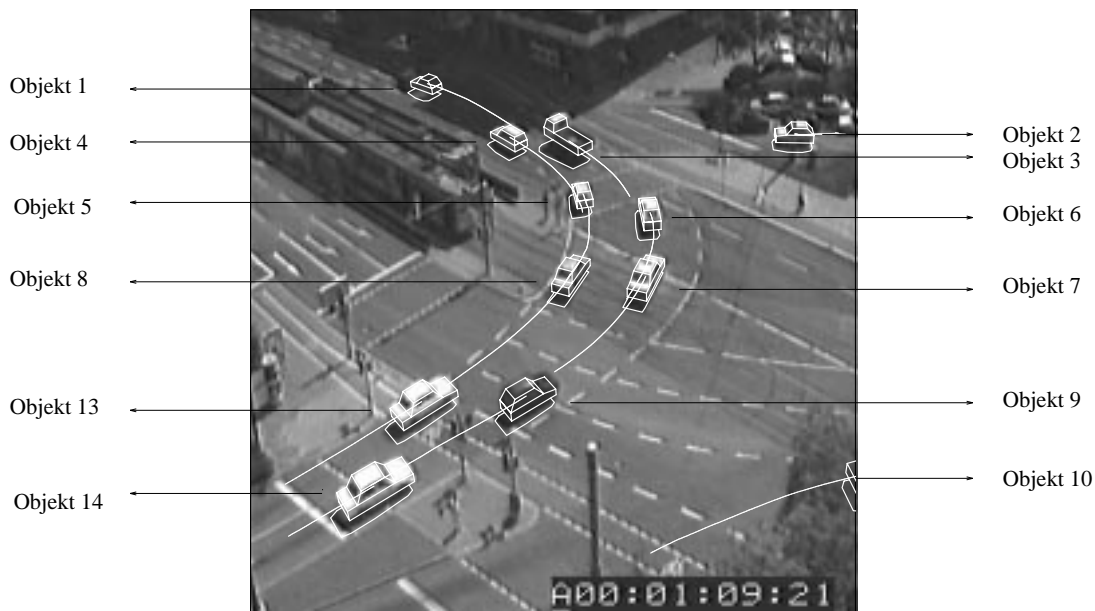


Abbildung A.11: Straßenverkehrsszene am Durlacher-Tor, zur Verdeutlichung mit eingezeichneten Trajektorien, wobei die automatisch erkannten Objekte durch umschreibende Linienmodelle eingegrenzt sind. Desweiteren ist zu jedem Fahrzeug die systemintern vergebene Objekt Nummer angegeben ([Kollnig 95]).

Abbildung A.11 zeigt ein Einzelbild der Durlacher-Tor-Szene. Bezüglich jedes erkannten Fahrzeuges dieser nur knapp zwei Sekunden langen Szene wurden natürlichsprachliche



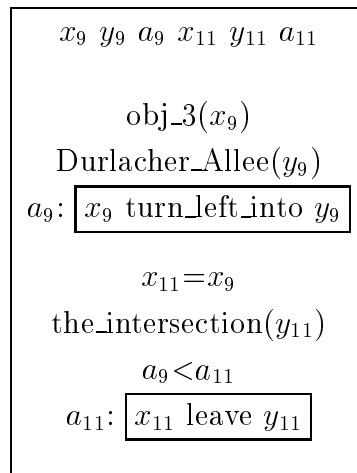
Beschreibungen seines Fahrverhaltens berechnet. Im folgenden ist ein Auswertungsergebnis für den Pritschenwagen ‘Objekt\_3’ der Szene angegeben. Für dieses Fahrzeug wurden folgende Situationen erkannt:

5 : 46 ! turning\_left\_into(obj\_3, durlacher\_allee).  
 47 : 95 ! leaving\_intersection(obj\_3).

Dieses Ergebnis wurde in folgenden natürlichsprachlichen Text umgesetzt:

Obj 3 turned left into Durlacher Allee. Then it left the intersection.

Abschließend ist dieses Ergebnis in seiner Darstellung als automatisch berechnete DRS wiedergegeben:



### A.3 Ettliger-Tor-Szene

Abbildung A.12 zeigt ein Einzelbild der Ettliger-Tor-Szene. Bezüglich jedes erkannten Fahrzeuges dieser nur knapp zwei Sekunden langen Szene wurden natürlichsprachliche Beschreibungen seines Fahrverhaltens berechnet. Im folgenden ist ein Auswertungsergebnis für das Fahrzeug ‘Objekt\_4’ der Szene angegeben, diesmal als szenische Beschreibung aus Protagonistensicht, wobei das Agens sein Verhalten aus seiner Sicht beschreibt:

I came from Kriegsstrasse. Then I turned left into Ettliger Strasse.

Hier ist die Situationsfolge, die bezüglich dieses Fahrzeuges berechnet wurde, und aus der der obige Text abgeleitet wurde:

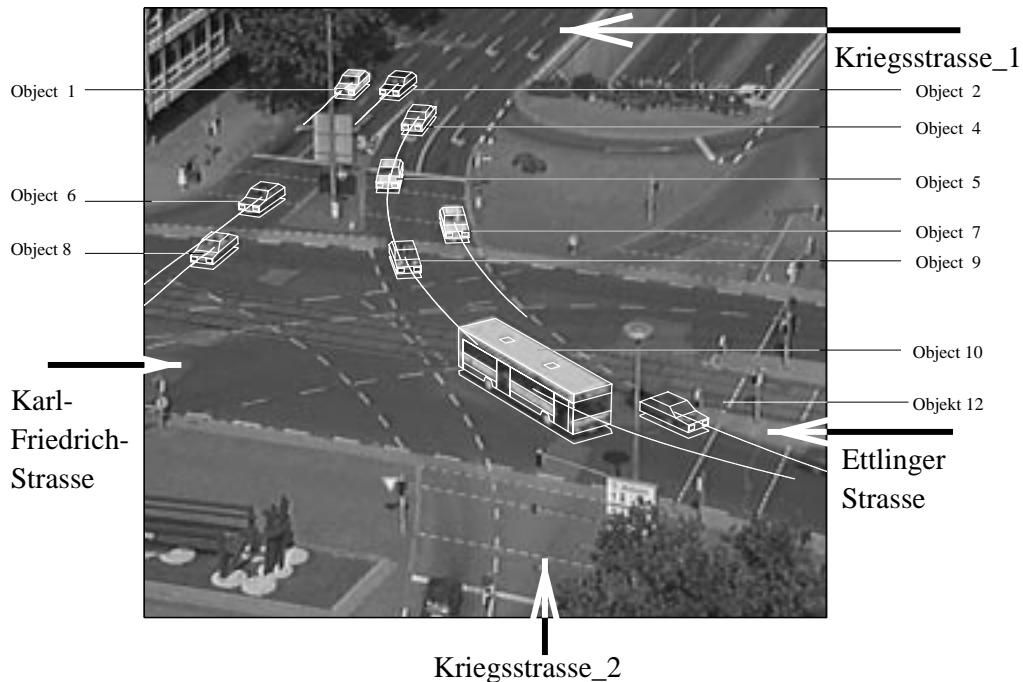


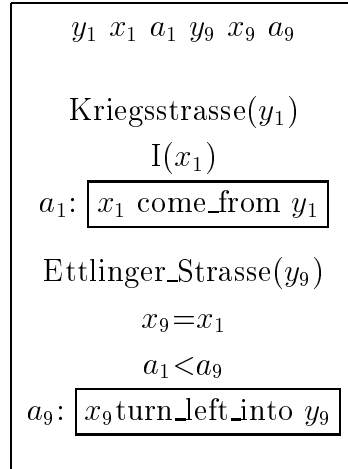
Abbildung A.12: Straßenverkehrsszene am Ettlinger-Tor, zur Verdeutlichung mit eingezeichneten Trajektorien, wobei die automatisch erkannten Objekte durch umschreibende Linienmodelle eingegrenzt sind. Desweiteren ist zu jedem Fahrzeug die systemintern vergebene Objekt Nummer angegeben ([Kollnig 95]).

```

5 : 5 ! coming_from(obj_4, kriegsstrasse).
6 : 68 ! taking_left_hand_turn_off_lane(obj_4).
69 : 95 ! turning_left_into(obj_4, ettlinger_strasse,).

```

Abschließend ist dieses Ergebnis in seiner Darstellung als automatisch berechnete DRS wiedergegeben. Dabei wurde zum zweiten Faktum bei der Perspektivierung keine Beschreibung assoziiert, da es als redundant betrachtet wurde. So findet es in der DRS und auch im Text keine Entsprechung. Da eine szenische Beschreibung aus Protagonistensicht abgeleitet werden sollte, wurde die Darstellung bei der Perspektivierung in 'Ich-Form' übertragen.



## A.4 Nibelungenplatzszene

Zu dieser Bildfolge wurde folgende quantorisierte Beschreibung berechnet, in der das gemeinsame Verhalten mehrerer Fahrzeuge in quantorisierte Aussagen umgesetzt wurde:

Eight vehicles came from road1. Four of them turned left into road2. Four vehicles drove straight ahead.

Erläuterung: Die Straßen wurden im Uhrzeigersinn mit ‘Road1’ bis ‘Road4’ bezeichnet. ‘Road1’ bezeichnet dabei diejenige Straße, die in Abbildung A.13 von oben zur Kreuzung führt. Das vorstehend angegebene Auswertungsergebnis ist unter Berücksichtigung folgender Einschränkungen korrekt: Die acht Fahrzeuge, von denen erkannt wurde, daß sie aus der Straße ‘Road1’ kamen, tragen folgende Objektnummern: 4, 5, 7, 8, 10, 11, 22, 23. Der menschliche Betrachter kann zwar erkennen, daß auch die Fahrzeuge 12, 14, 15, 16, 17, 18, 19 aus dieser Straße gekommen sein dürften. Entsprechende Aussage wurde vom System jedoch nicht generiert, da sich letztgenannte Fahrzeuge bereits zu Beginn der Auswertung in der Kreuzungsmitte befanden (siehe Abbildung A.13, welches das **erste** Bild der Auswertung zeigt). Zu Auswertungsbeginn befanden sich also nur erstgenannte Fahrzeuge noch in dem Zufahrtsbereich der Kreuzung, der der Straße ‘Road1’ zugeordnet wird. Für die Fahrzeuge, die sich schon zu Beginn der Kreuzung in der ‘neutralen’ Kreuzungsmitte befanden, wird hingegen nicht mehr über deren Herkunft ‘spekuliert’.



Abbildung A.13: Erstes Einzelbild einer Bildfolge, die am Nibelungenplatz in Frankfurt aufgezeichnet wurde. Zu jedem verfolgten Fahrzeug ist die systemintern vergebene Objekt Nummer angegeben. Zur Beibehaltung der Übersichtlichkeit wurden die berechneten Trajektorien nicht eingezeichnet.

## A.5 Rheinhafenszene

Die Rheinhafenszene unterscheidet sich von allen anderen Kreuzungsszenen dadurch, daß sie keine Ampelsteuerung besitzt. Dies bedeutet, daß es hier keine eindeutig selektierbaren Verkehrsströme gibt. Stattdessen ist zu erwarten, daß der Strom der vorfahrtsberechtigten Fahrzeuge aus der Honsellstraße (siehe Abbildung A.14) zwischenzeitlich von einzelnen Fahrzeugen unterbrochen wird, die sich langsam von nicht vorfahrtsberechtigenden Kreuzungsarmen kommend in die Kreuzung einflechten. Dies spiegelt sich auch in den Beschreibungen wieder, die für die Fahrzeuge berechnet wurden (hier eine Auswahl).



Abbildung A.14: Straßenverkehrsszene am Rheinhafen mit fünf automatisch berechneten Fahrzeugtrajektorien und Modellüberlagerungen. Rechts die *Wikingerstraße*, unten die *Honselstraße* und links die *Rheinhafenstraße*.

Obj 2 came from Honselstrasse. Then it turned left into Rheinhafenstrasse.

Obj 5 came from Honselstrasse. Then it turned right into Wikingerstrasse. Then it left the intersection.

Obj 4 came from Wikingerstrasse. Then it stopped in order to give way to other vehicles which passed the intersection.

Zu letztgenanntem Fahrzeug ist die korrespondierende automatisch berechnete begriffliche Beschreibung wie folgt angegeben:

2139 : 2139 ! come\_from(obj\_4, wikingerstrasse, 1).  
2140 : 2140 ! stop(obj\_4, 2).  
2140 : 2140 ! give\_way\_to(obj\_4, other\_vehicles, 3).  
2140 : 2140 ! reason(4, 2, 3).  
2140 : 2140 ! pass(other\_vehicles, the\_intersection, 6).

# Anhang B

## Situationsgraphenbaum

Im folgenden ist der in dieser Arbeit entwickelte und verwendete Situationsbestimmungsbaum zur Ableitung natürlichsprachlicher Beschreibung von Straßenverkehrsszenen an Tankstellen angegeben. Der Aufbau des Baumes ist in Abschnitt 3.4.2 natürlichsprachlich beschrieben. Es folgen einige Ausführungen zur Notation.

Jeder Graph des Situationsgraphenbaums wird durch das Schlüsselwort ‘GRAPH’ eingeleitet. Es folgt der Name des Graphen, ggf. mit Angabe einiger zusätzlicher Parameter. Die einzelnen Situationsgraphen eines Situationsgraphenbaums sind hierarchisch angeordnet. Deshalb folgt dem Namen des Graphen, abgetrennt durch einen Doppelpunkt, der Name der Vatersituation. Im vorliegenden Fall besitzt der Graph ‘faktenberechnung’ keine Vatersituation, bildet also somit die Wurzel des Graphenbaums. Der Graph ‘gr\_aktiv’ hingegen besitzt die Situation ‘sit\_aktiv’ als Vater. Ein Graph besteht aus einer oder mehreren Situationen. Die Situationsbeschreibungen werden durch das Schlüsselwort ‘SIT’ eingeleitet. Ihm folgt der Situationsname, wiederum gefolgt durch die Aufzählung potentieller Nachfolgesituationen. Es folgen, jeweils abgegrenzt durch geschweifte Klammern, die Angabe des Zustands- und des Handlungsschemas der Situation (vgl. Abschnitt 3.3). Das im Handlungsschema oft verwendete Prädikat ‘note’ erlaubt den Eintrag der darin enthaltenen Beschreibung in die Faktendatenbank.

Einzelheiten über die automatische Transformation eines solchen Graphenbaums in ein unscharf metrisch temporales Logikprogramm entnehme man [Schäfer 96] oder [Haag&Nagel 98c].

```

/*****
// Situationsgraphenbaum "Kreuz1" - Version 1    ... (wt)
DEFAULT NONINCREMENTAL GREEDY PLURAL DEPTH; // INTERACTIVE;
/*****
GRAPH faktenberechnung(Agens)
{
  SIT sit_append
    : sit_inaktiv,
      sit_aktiv
  {
  }
  {
  append('ergebnis.lim');
  }

  SIT sit_aktiv
  {
    agens(Agens);
    traj_aktiv(Agens);
  }
  { }

  SIT sit_inaktiv
    : sit_inaktiv,
      sit_aktiv
  {
    agens(Agens);
    inaktiv(Agens);
  }
  {
  }
}

/*****
GRAPH gr_aktiv : sit_aktiv
{
  SIT sit_getting_petrol
    : sit_getting_petrol
  {
    tanksaeule(TP_beliebig);
  }
  {
    note(getting_petrol(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_crossing_intersection1
    : sit_crossing_intersection1
  {
    knotenpunktzentrum(KPZ_beliebig);
  }
  {
    note(crossing_intersection1(Agens));
    note(traj_aktiv(Agens));
  }
}

GRAPH gr_getting_petrol : sit_getting_petrol
{
  SIT sit_driving_to_dispensing_pump
    : sit_driving_to_dispensing_pump(Spur),
      sit_driving_to_dispensing_pump,
      sit_taking_in_petrol(Spur)
  {
    tank_ausweich_spur(Spur);
    noch_kein_tp_erreicht(Agens);
  }
  {
    note(driving_to_dispensing_pump(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_taking_in_petrol
    : sit_taking_in_petrol,
      sit_leaving_dispensing_pump
  {
    auf(Agens, tankplatz: Tankplatz);
    v_Betrag(Agens, null) AT 0.3;
  }
  {
    note(taking_in_petrol(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_leaving_dispensing_pump
    : sit_leaving_dispensing_pump,
      sit_leaving_filling_station
  {
    diessseitiger_aufenthalt(Agens);
  }
  {
    note(leaving_dispensing_pump(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_leaving_filling_station
    : sit_leaving_filling_station
  {
    aufenthalt(Agens, abseits);
  }
  {
    note(leaving_filling_station(Agens));
    note(traj_aktiv(Agens));
  }
}

GRAPH gr_driving_to_dispensing_pump : sit_driving_to_dispensing_pump
{
  SIT sit_wishing_to_tank
    : sit_driving_on_filling_lane
  {
    tankspur(Spur);
  }
  {
    note(wishing_to_tank(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_driving_on_filling_lane
    : sit_driving_on_filling_lane
  {
    tank_ausweich_spur(Spur);
  }
  {
    note(driving_on_filling_lane(Agens));
    note(traj_aktiv(Agens));
  }
}

GRAPH gr_driving_on_filling_lane : sit_driving_on_filling_lane
{
  SIT sit_filling_place_occupation_lower_lane
    : sit_no_filling_place_occupied_on_lower_filling_lane,
      sit_one_filling_place_occupied_on_lower_filling_lane,
      sit_two_filling_places_occupied_on_lower_filling_lane,
      sit_all_filling_places_occupied_on_lower_filling_lane
  {
    v_tankplatz(Tplatz1);
    name_tankplatz(Tplatz1, the_last_filling_place);
    v_tankplatz(Tplatz2);
    name_tankplatz(Tplatz2, the_first_filling_place);
    v_tankplatz(Tplatz3);
    name_tankplatz(Tplatz3, the_second_filling_place);
  }
  {
    note(filling_place_occupation_lower_lane(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_no_filling_place_occupied_on_lower_filling_lane
    : sit_filling_place_occupation_upper_lane(Tplatz1, Tplatz2, Tplatz3)
  {
    null_belegt(Tplatz1, Tplatz2, Tplatz3);
  }
  {
    note(no_filling_place_occupied_on_lower_filling_lane);
    note(traj_aktiv(Agens));
  }

  SIT sit_one_filling_place_occupied_on_lower_filling_lane
    : sit_filling_place_occupation_upper_lane(Tplatz1, Tplatz2, Tplatz3)
  {
    einer_belegt(Tplatz1, Tplatz2, Tplatz3);
  }
  {
    note(one_filling_place_occupied_on_lower_filling_lane);
    note(traj_aktiv(Agens));
  }

  SIT sit_two_filling_places_occupied_on_lower_filling_lane
    : sit_filling_place_occupation_upper_lane(Tplatz1, Tplatz2, Tplatz3)
  {
    zwei_belegt(Tplatz1, Tplatz2, Tplatz3);
  }
  {
    note(two_filling_places_occupied_on_lower_filling_lane);
    note(traj_aktiv(Agens));
  }
}

```



```

SIT sit_all_filling_places_occupied_on_lower_filling_lane
: sit_filling_place_occupation_upper_lane(Tplatz1,Tplatz2,Tplatz3)
{
  drei_belegt(Tplatz1,Tplatz2,Tplatz3);
}
{
  note(all_filling_places_occupied_on_lower_filling_lane);
  note(traj_aktiv(Agens));
}

SIT sit_filling_place_occupation_upper_lane
: sit_no_filling_place_occupied_on_upper_filling_lane,
  sit_one_filling_place_occupied_on_upper_filling_lane,
  sit_two_filling_places_occupied_on_upper_filling_lane,
  sit_all_filling_places_occupied_on_upper_filling_lane
{
  h_tankplatz(Tplatz1);
  name_tankplatz(Tplatz1,the_last_filling_place);
  h_tankplatz(Tplatz2);
  name_tankplatz(Tplatz2,the_first_filling_place);
  h_tankplatz(Tplatz3);
  name_tankplatz(Tplatz3,the_second_filling_place);
}
{
  note(filling_place_occupation_upper_lane(Agens));
  note(traj_aktiv(Agens));
}

SIT sit_no_filling_place_occupied_on_upper_filling_lane
: sit_choosing_lower_filling_lane,
  sit_choosing_upper_filling_lane
{
  null_belegt(Tplatz1,Tplatz2,Tplatz3);
}
{
  note(no_filling_place_occupied_on_upper_filling_lane);
  note(traj_aktiv(Agens));
}

SIT sit_one_filling_place_occupied_on_upper_filling_lane
: sit_choosing_lower_filling_lane,
  sit_choosing_upper_filling_lane
{
  einer_belegt(Tplatz1,Tplatz2,Tplatz3);
}
{
  note(one_filling_place_occupied_on_upper_filling_lane);
  note(traj_aktiv(Agens));
}

SIT sit_two_filling_places_occupied_on_upper_filling_lane
: sit_choosing_lower_filling_lane,
  sit_choosing_upper_filling_lane
{
  zwei_belegt(Tplatz1,Tplatz2,Tplatz3);
}
{
  note(two_filling_places_occupied_on_upper_filling_lane);
  note(traj_aktiv(Agens));
}

SIT sit_all_filling_places_occupied_on_upper_filling_lane
: sit_choosing_lower_filling_lane,
  sit_choosing_upper_filling_lane
{
  drei_belegt(Tplatz1,Tplatz2,Tplatz3);
}
{
  note(all_filling_places_occupied_on_upper_filling_lane);
  note(traj_aktiv(Agens));
}

SIT sit_choosing_lower_filling_lane
: sit_choosing_dispensing_pump
{
  auf(Agens,Spur7);
  u_tankspur(Spur7);
}
{
  note(choosing_lower_filling_lane(Agens));
  note(traj_aktiv(Agens));
}

SIT sit_choosing_upper_filling_lane
: sit_choosing_dispensing_pump
{
  auf(Agens,Spur7);
  o_tankspur(Spur7);
}
}

}

SIT sit_choosing_dispensing_pump
: sit_choosing_dispensing_pump
{
  o_tank_ausweichspur(Spur);
}
{
  note(traj_aktiv(Agens));
  note(choosing_dispensing_pump(Agens));
}
}

GRAPH gr_choosing_dispensing_pump : sit_choosing_dispensing_pump
{
  SIT sit_first_filling_place_occupied
  : sit_approach_occupied_first_filling_place,
  sit_overtaking_vehicle
  {
    tankplatz(T1o);
    name_tankplatz(T1o,the_first_filling_place);
    tankplatz_belegt(T1o);
  }
  {
    note(first_filling_place_occupied);
    note(traj_aktiv(Agens));
  }

  SIT sit_second_filling_place_occupied
  : sit_driving_to_first_filling_place,
  sit_overtaking_vehicle
  {
    tankplatz(T4);
    name_tankplatz(T4,the_second_filling_place);
    tankplatz_belegt(T4);
  }
  {
    note(second_filling_place_occupied);
    note(traj_aktiv(Agens));
  }

  SIT sit_driving_to_first_filling_place
  : sit_driving_to_first_filling_place,
  sit_overtaking_vehicle,
  sit_reaching_first_filling_place
  {
    hinfahrend(Agens);
    auf(Agens, tankspur: Spur1);
  }
  {
    note(driving_to_first_filling_place(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_reaching_first_filling_place
  : sit_reaching_first_filling_place,
  sit_driving_to_second_filling_place
  {
    fahrend(Agens);
    tankplatz(T1o);
    auf(Agens,T1o);
  }
  {
    note(reaching_first_filling_place(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_driving_to_second_filling_place
  : sit_driving_to_second_filling_place,
  sit_reaching_second_filling_place
  {
    hinfahrend(Agens);
  }
  {
    note(driving_to_second_filling_place(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_reaching_second_filling_place
  : sit_reaching_second_filling_place,
  sit_driving_to_last_filling_place
  {
    fahrend(Agens);
  }
}

```

```

    tankplatz(T2o);
    name_tankplatz(T2o,the_second_filling_place);
    auf(Agens,T2o);
}
{
    note(reaching_second_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_driving_to_last_filling_place
: sit_driving_to_last_filling_place,
  sit_reaching_last_filling_place
{
    hinfahrend(Agens);
}
{
    note(driving_to_last_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_reaching_last_filling_place
: sit_reaching_last_filling_place,
  sit_passing_last_filling_place
{
    fahrend(Agens);
    tankplatz(T3o);
    name_tankplatz(T3o,the_last_filling_place);
    auf(Agens,T3o);
}
{
    note(reaching_last_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_overtaking_vehicle
: sit_overtaking_vehicle,
  sit_change_back_to_filling_lane
{
    auf(Agens,ausweichspur: Spur2);
}
{
    note(overtaking_vehicle(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_change_back_to_filling_lane
: sit_reaching_second_filling_place,
  sit_reaching_last_filling_place,
  sit_driving_to_next_dispensing_pump
{
    auf(Agens,tankspur: Spur4);
}
{
    note(change_back_to_filling_lane(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_driving_to_next_dispensing_pump
: sit_driving_to_next_dispensing_pump,
  sit_reaching_last_filling_place,
  sit_remain_standing_behind_last_filling_place,
  sit_backing_up_behind_last_filling_place
{
    hinfahrend(Agens);
}
{
    note(driving_to_next_dispensing_pump(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_remain_standing_behind_last_filling_place
: sit_remain_standing_behind_last_filling_place,
  sit_backing_up_behind_last_filling_place,
  sit_passing_last_filling_place
{
    v_Betrag(Agens,null) AT 0.1;
}
{
    note(remain_standing_behind_last_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_backing_up_behind_last_filling_place
: sit_backing_up_behind_last_filling_place,
  sit_reaching_last_filling_place,
  sit_remain_standing_behind_last_filling_place
{
    modus(Agens, rueckwaerts) AT 0.1;
}

    not_auf_tp(Agens);
}
{
    note(backing_up_behind_last_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_approach_occupied_first_filling_place
: sit_approach_occupied_first_filling_place,
  sit_overtaking_vehicle,
  sit_waiting_in_front_of_first_filling_place,
  sit_reaching_first_filling_place
{
    hinfahrend(Agens);
    auf(Agens,Spur5);
    tankspur(Spur5);
}
{
    note(approach_occupied_first_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_waiting_in_front_of_first_filling_place
: sit_waiting_in_front_of_first_filling_place,
  sit_approach_occupied_first_filling_place
{
    v_Betrag(Agens,null);
}
{
    note(waiting_in_front_of_first_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_passing_last_filling_place
: sit_passing_last_filling_place,
  sit_remain_standing_behind_last_filling_place,
  sit_backing_up_behind_last_filling_place
{
    hinfahrend(Agens);
}
{
    note(passing_last_filling_place(Agens));
    note(traj_aktiv(Agens));
}

}

GRAPH gr_leaving_dispensing_pump : sit_leaving_dispensing_pump
{
    SIT sit_leaving_first_filling_place
: sit_second_filling_place_occupied2,
  sit_last_filling_place_occupied,
  sit_free_exit
{
    auf(Agens,T1v);
    tankplatz(T1v);
    name_tankplatz(T1v,the_first_filling_place);
}
{
    note(leaving_first_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_leaving_second_filling_place
: sit_last_filling_place_occupied,
  sit_free_exit
{
    auf(Agens,T2v);
    tankplatz(T2v);
    name_tankplatz(T2v,the_second_filling_place);
}
{
    note(leaving_second_filling_place(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_free_exit
: sit_free_exit
{
    diessseitiger_aufenthalt(Agens);
}
{
    note(free_exit(Agens));
    note(traj_aktiv(Agens));
}

SIT sit_second_filling_place_occupied2
: sit_no_free_exit

```

```

{
  auf(Agens,T1v);
  aufbringung(Spur_ou,T1v);
  aufbringung(Spur_ou,T2v);
  name_tankplatz(T2v,the_second_filling_place);
  tankplatz_belegt(T2v);
}
{ note(second_filling_place_occupied);
  note(traj_aktiv(Agens));
}

SIT sit_last_filling_place_occupied
: sit_no_free_exit
{
  auf(Agens,T1v);
  aufbringung(Spur_ou,T1v);
  aufbringung(Spur_ou,T3v);
  name_tankplatz(T3v,the_last_filling_place);
  tankplatz_belegt(T3v);
}
{ note(last_filling_place_occupied);
  note(traj_aktiv(Agens));
}

SIT sit_no_free_exit
{
  diesseitiger_aufenthalt(Agens);
}
{
  note(no_free_exit(Agens));
  note(traj_aktiv(Agens));
}
}

GRAPH gr_no_free_exit : sit_no_free_exit
{
  SIT sit_approaching_patiens_when_leaving
  : sit_approaching_patiens_when_leaving,
  sit_waiting_behind_patiens_when_leaving,
  sit_passing_patiens_when_leaving
  {
    fahrend(Agens);
    modus(Agens, vorwaerts);
    auf(Agens, tankspur: Spur_x);
  }
  { note(traj_aktiv(Agens));
  note(approaching_patiens_when_leaving(Agens)); }

  SIT sit_backing_up_behind_patiens_when_leaving
  : sit_backing_up_behind_patiens_when_leaving,
  sit_waiting_behind_patiens_when_leaving
  {
    fahrend(Agens);
    modus(Agens, rueckwaerts);
    auf(Agens, tankspur: Spur_x);
  }
  { note(traj_aktiv(Agens));
  note(backing_up_behind_patiens_when_leaving(Agens)); }

  SIT sit_waiting_behind_patiens_when_leaving
  : sit_waiting_behind_patiens_when_leaving,
  sit_driving_off_behind_patiens_when_leaving,
  sit_passing_patiens_when_leaving,
  sit_backing_up_behind_patiens_when_leaving
  {
    v_Betrag(Agens,null) AT 0.1;
  }
  { note(traj_aktiv(Agens));
  note(waiting_behind_patiens_when_leaving(Agens)); }

  SIT sit_driving_off_behind_patiens_when_leaving
  : sit_driving_off_behind_patiens_when_leaving,
  sit_passing_patiens_when_leaving
  {
    fahrend(Agens);
    modus(Agens, vorwaerts);
    auf(Agens, tankspur: Spur_x);
  }
  { note(traj_aktiv(Agens));
  note(driving_off_behind_patiens_when_leaving(Agens)); }

  SIT sit_passing_patiens_when_leaving
  : sit_passing_patiens_when_leaving
  {
    fahrend(Agens);
    modus(Agens, vorwaerts);
    auf(Agens, ausweichspur: Spur_y);
  }
}

{
  note(traj_aktiv(Agens));
  note(passing_patiens_when_leaving(Agens)); }
}

GRAPH gr_crossing_intersection1 : sit_crossing_intersection1
{
  SIT sit_provenance
  : sit_driving_to_intersection
  {
    auf(Agens, kreuzungszufahrt : Bahn);
    f_richtung(Agens, Bahn, laengs);
  }
  { note(traj_aktiv(Agens));
  note(provenance(Agens)); }

  SIT sit_driving_to_intersection
  : sit_driving_to_intersection,
  sit_crossing_intersection(Bahn)
  {
    auf(Agens, kreuzungszufahrt : Bahn);
    f_richtung(Agens, Bahn, laengs);
  }
  { note(traj_aktiv(Agens));
  note(driving_to_intersection(Agens)); }

  SIT sit_crossing_intersection
  : sit_crossing_intersection,
  sit_crossing_intersection(Bahn),
  sit_leaving_intersection(Bahn)
  {
    auf(Agens, kreuzungsweg : Bahn);
    f_richtung(Agens, Bahn, laengs);
  }
  { note(traj_aktiv(Agens));
  note(crossing_intersection(Agens)); }

  SIT sit_leaving_intersection
  : sit_leaving_intersection,
  sit_leaving_intersection(Bahn)
  {
    auf(Agens, kreuzungsabfahrt : Bahn);
    f_richtung(Agens, Bahn, laengs);
  }
  { note(traj_aktiv(Agens));
  note(leaving_intersection(Agens)); }
}

GRAPH gr_provenance : sit_provenance
{
  SIT sit_coming_from
  {
    fahrbahnname(Bahn, Name);
  }
  { note(traj_aktiv(Agens));
  note(coming_from(Agens,Name)); }
}

GRAPH gr_driving_to_intersection : sit_driving_to_intersection
{
  SIT sit_approach_intersection
  : sit_approach_intersection,
  sit_stopping_at_intersection
  {
    v_Betrag(Agens, g1kl) AT 0.1;
  }
  { note(traj_aktiv(Agens));
  note(approach_intersection(Agens)); }

  SIT sit_stopping_at_intersection
  : sit_waiting_at_intersection,
  sit_approach_intersection
  {
    v_Betrag(Agens, null) AT 0.1;
  }
  { note(traj_aktiv(Agens));
  note(stopping_at_intersection(Agens)); }

  SIT sit_waiting_at_intersection
  : sit_waiting_at_intersection,
  sit_starting_at_intersection
  {
    v_Betrag(Agens, null) AT 0.1;
  }
  { note(traj_aktiv(Agens));
  note(waiting_at_intersection(Agens)); }

  SIT sit_starting_at_intersection

```

```

: sit_starting_at_intersection,
  sit_stopping_at_intersection
{
  v_Betrag(Agens, glk1) AT 0.1;
}
{
  note(traj_aktiv(Agens));
}
note(starting_at_intersection(Agens)); }
}

GRAPH gr_approach_intersection : sit_approach_intersection
{
  SIT sit_taking_left_hand_turn_off_lane
  : sit_taking_left_hand_turn_off_lane,
  sit_taking_straight_ahead_lane
  {
    kr_zufahrt_links(Bahn);
  }
  {
    note(taking_left_hand_turn_off_lane(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_taking_straight_ahead_lane
  : sit_taking_straight_ahead_lane,
  sit_taking_right_hand_turn_off_lane,
  sit_taking_left_hand_turn_off_lane
  {
    kr_zufahrt_gerade(Bahn);
  }
  {
    note(taking_straight_ahead_lane(Agens));
    note(traj_aktiv(Agens));
  }

  SIT sit_taking_right_hand_turn_off_lane
  : sit_taking_right_hand_turn_off_lane,
  sit_taking_straight_ahead_lane
  {
    kr_zufahrt_rechts(Bahn);
  }
  {
    note(taking_right_hand_turn_off_lane(Agens));
  }
}

GRAPH gr_crossing_intersection : sit_crossing_intersection
{
  SIT sit_turning_right
  : sit_turning_right,
  sit_turning_right(Bahn, Bahn2)
  {
    vor(Bahn2, Bahn);
    kr_zufahrt_rechts(Bahn2);
  }
  {
    note(traj_aktiv(Agens));
  }
  note(turning_right(Agens)); }

  SIT sit_driving_straight_ahead
  : sit_driving_straight_ahead,
  sit_driving_straight_ahead(Bahn, Bahn2)
  {
    vor(Bahn2, Bahn);
    kr_zufahrt_gerade(Bahn2);
  }
  {
    note(traj_aktiv(Agens));
  }
  note(driving_straight_ahead(Agens)); }

  SIT sit_turning_left
  : sit_turning_left,
  sit_turning_left(Bahn, Bahn2)
  {
    vor(Bahn2, Bahn);
    kr_zufahrt_links(Bahn2);
  }
  {
    note(traj_aktiv(Agens));
  }
  note(turning_left(Agens)); }
}

/*****
GRAPH gr_turning_left : sit_turning_left
{
  SIT sit_turning_left_into
  : sit_turning_left_into(Bahn)
  {
    hinfuehrend_zu(Bahn, Name2);
  }
  {
    note(traj_aktiv(Agens));
  }
}

note(turning_left_into(Agens, Name2)); }
}

/*****
GRAPH gr_turning_right : sit_turning_right
{
  SIT sit_turning_right_into
  : sit_turning_right_into(Bahn)
  {
    hinfuehrend_zu(Bahn, Name2);
  }
  {
    note(traj_aktiv(Agens));
  }
  note(turning_right_into(Agens, Name2)); }
}

/*****
TERM nicht_auf(Agens, Spur)
{
  NOT { auf(Agens, Spur); }
}

TERM spurwechseln(Agens, Spur, Spur2)
{
  AND { ausweichspur(Spur2);
        auf(Agens, Spur2);
        Spur <> Spur2;
      }
}

TERM null_belegt(Tplatz1, Tplatz2, Tplatz3)
{
  AND { tankplatz_frei(Tplatz1);
        tankplatz_frei(Tplatz2);
        tankplatz_frei(Tplatz3);
      }
}

TERM einer_belegt(Tplatz1, Tplatz2, Tplatz3)
{
  OR { frei_frei_belegt(Tplatz1, Tplatz2, Tplatz3);
       belegt_frei_frei(Tplatz1, Tplatz2, Tplatz3);
       frei_belegt_frei(Tplatz1, Tplatz2, Tplatz3);
     }
}

TERM frei_frei_belegt(Tplatz1, Tplatz2, Tplatz3)
{
  AND { tankplatz_frei(Tplatz1);
        tankplatz_frei(Tplatz2);
        tankplatz_belegt(Tplatz3);
      }
}

TERM belegt_frei_frei(Tplatz1, Tplatz2, Tplatz3)
{
  AND { tankplatz_belegt(Tplatz1);
        tankplatz_frei(Tplatz2);
        tankplatz_frei(Tplatz3);
      }
}

TERM frei_belegt_frei(Tplatz1, Tplatz2, Tplatz3)
{
  AND { tankplatz_frei(Tplatz1);
        tankplatz_belegt(Tplatz2);
        tankplatz_frei(Tplatz3);
      }
}

TERM zwei_belegt(Tplatz1, Tplatz2, Tplatz3)
{
  OR { belegt_frei_belegt(Tplatz1, Tplatz2, Tplatz3);
       belegt_belegt_frei(Tplatz1, Tplatz2, Tplatz3);
       frei_belegt_belegt(Tplatz1, Tplatz2, Tplatz3);
     }
}

TERM belegt_frei_belegt(Tplatz1, Tplatz2, Tplatz3)
{
  AND { tankplatz_belegt(Tplatz1);
        tankplatz_frei(Tplatz2);
        tankplatz_belegt(Tplatz3);
      }
}

TERM belegt_belegt_frei(Tplatz1, Tplatz2, Tplatz3)
{
}

```

```

    AND { tankplatz_belegt (Tplatz1);
          tankplatz_belegt (Tplatz2);
          tankplatz_frei (Tplatz3);
        }
  }

TERM frei_belegt_belegt (Tplatz1,Tplatz2,Tplatz3)
{
  AND { tankplatz_frei (Tplatz1);
        tankplatz_belegt (Tplatz2);
        tankplatz_belegt (Tplatz3);
      }
}

TERM drei_belegt (Tplatz1,Tplatz2,Tplatz3)
{
  AND { tankplatz_belegt (Tplatz1);
        tankplatz_belegt (Tplatz2);
        tankplatz_belegt (Tplatz3);
      }
}

TERM tankplatz_belegt (Tplatz)
{
  AND { patients (Fzg);
        auf (Fzg,Tplatz);
      }
}

TERM tankplatz_frei (Tplatz)
{
  NOT { tankplatz_belegt (Tplatz);
        }
}

TERM inaktiv (Agens)
{
  NOT { traj_aktiv (Agens); }
}

TERM fahrend (Kfz) {
  OR { v_Betrag (Kfz, sehr_klein);
        v_Betrag (Kfz, glkl);
      }
}

TERM tank_abweich_spur (Spur) {
  OR { tankspur (Spur);
        ausweichspur (Spur); }
}

TERM hinfahrend (Kfz) {
  AND { fahrend (Kfz);
        not_auf_tp (Kfz); }
}

TERM noch_kein_tp_erreicht (Kfz) {
  OR { fahrend (Kfz);
        not_auf_tp (Kfz); }
}

TERM v_Betrag_nicht_glkl (Agens) {
  NOT { v_Betrag (Agens, glkl) AT 1.99; }
}

TERM not_auf_tp (Kfz) {
  NOT { auf_tp (Kfz); }
}

TERM auf_tp (Kfz) {
  auf (Kfz, tankplatz:X);
}

TERM o_tank_abweichspur (Spur) {
  OR { tankspur (Spur);
        ausweichspur (Spur); }
}

TERM diesseitiger_aufenthalt (Agens) {
  NOT { aufenthalt (Agens, abseits); }
}

```

# Anhang C

## MALCOLM

Das Programm MALCOLM erstellt aus einem gegebenen Situationsgraphen (Situationsbestimmungsbaum) und einer gegebenen Perspektive automatisch einen weiteren Situationsgraphenbaum (Perspektivierungsbaum), der die Ableitung einer perspektivischen Beschreibung mit der gegebenen Perspektive aus der durch den Situationsbestimmungsbaum ermittelten Situationsfolge für ein Agens erlaubt. Nähere Informationen über die Wirkungsweise von MALCOLM finden sich in Abschnitt 3.4.4. Der Perspektivierungsbaum ist ein zweistufiger Situationsgraphenbaum. Sein Wurzelgraph ist eine Kopie des Wurzelgraphen des zugehörigen Situationsbestimmungsbaumes. Sein Untergraph enthält alle Situationsbeschreibungen, die nach oben angegebener Vorgehensweise aus dem Situationsbestimmungsbaum und der Perspektivierungstabelle abgeleitet wurden.

Im folgenden sind Entwurfs- und Benutzungsinformationen über MALCOLM aufgeführt. **Fettgedruckte** Wörter und Wortteile geben feststehende Schlüsselwörter an, während *kursive* Wortteile das Einsetzen freier Bezeichnungen erlauben. Das Zeichen  $\square$  bezeichnet zu berücksichtigende Leerzeichen.

MALCOLM wird folgendermaßen aufgerufen:

```
malcolm sitgraphname.sit perspektive.persp
```

### C.1 Formatvorgaben

Die Datei '*sitgraphname.sit*' enthält einen Situationsgraphen (Situationsbestimmungsbaum) in SIT++-Notation. Der Aufbau einzelner Graphen in der Datei muß darüberhinaus folgende Formatvorgaben erfüllen:

```
GRAPH gr_graphname : sit_situationsnamex  
{
```

```

□ □ □SIT sit_situationsnamex1
□ □ □: sit_nachfolgesituation1,
□ □ □ □sit_nachfolgesituation2,
           ∫
□ □ □ □sit_nachfolgesituationn
□ □ □{
□ □ □Praedikate des Zustandsschemas
□ □ □}
□ □ □{
□ □ □note(traj_aktiv(Agens));
□ □ □note(situationsnamex1(Beteiligte Individuen));
□ □ □}

□ □ □SIT sit_situationsnamex2
usw.
}

```

Die Datei 'perspektive.persp' enthält Angaben darüber, welche Fakten aus bestimmten Situationen des Situationsbestimmungsbaums bzw. aus bestimmten Situationsübergängen abgeleitet werden sollen. Die Datei muß folgende Formatvorgaben erfüllen:

1. Für perspektivische Situationsbeschreibungen:

```

sit_situationsname1
1: Faktum11
1: Faktum12
  ∫
1: Faktum1k
2: Faktum21
2: Faktum22
  ∫
2: Faktum2m
3: Faktum31
3: Faktum32
  ∫
3: Faktum3n
-----
sit_situationsname2
usw.

```

Die in der Datei 'perspektive.persp' verwendeten Situationsnamen sollten zweck-

mäßigerweise den Situationsnamen aus dem Situationsgraphen der Datei ‘*sit-graphname.sit*’ entsprechen. Es gilt außerdem:  $k \geq 0, m \geq 0, n \geq 0, m > 0 \rightarrow k > 0, n > 0 \rightarrow m > 0$ . Fakten gleicher Ebene  $eb \in \{1, 2, 3\}$  werden derselben Situation zugeordnet. Fakten der Ebene  $eb_2 > eb_1$  erhalten  $eb_2 - eb_1$  Auswertungszeitpunkte nach den Fakten der Ebene  $eb_1$  Gültigkeit.

2. Für perspektive Beschreibungen von Situationsübergängen:

**sit\_situationname<sub>i</sub>** **2** **sit\_situationname<sub>j</sub>**

**1:** *Faktum<sub>11</sub>*

**1:** *Faktum<sub>12</sub>*

*f*

**1:** *Faktum<sub>1k</sub>*

---

Es gilt  $i \neq j$  und  $k \geq 1$ . Situationsübergänge können nur auf der Ebene **1** beschrieben werden.

Das Lesen der beiden Dateien wird in MALCOLM von den Funktionen ‘**lese\_sitgraph**’ bzw. ‘**lese\_perspektivierung**’ und deren Unterfunktionen übernommen.

## C.2 Datenstrukturen

Nach dem Lesen der Datei, die den Situationsbestimmungsbaum enthält, werden dessen einzelne Situationen in einer Liste verwaltet. Jedes Glied der Liste repräsentiert genau eine Situation. Die Glieder haben den in Abbildung C.1 gezeigten Aufbau. Die Funktion ‘**umsortieren**’ sortiert die Situationen in eine Reihenfolge, die der linksbündigen (Präfix-)Traversierung des Situationsgraphenbaumes entspricht. Das erste Glied der Liste repräsentiert die erste Situation des Wurzelgraphen. Falls zu dieser Situation ein Untergraph mit spezielleren Situationen existiert, bildet die erste Situation dieses Untergraphen das zweite Glied der Liste. Dieser hierarchische Abstieg wird fortgesetzt, bis die erste Situation des speziellsten Untergraphen erreicht wurde. Dann wird mit der zweiten Situation des speziellsten Untergraphen analog fortgefahren. Sind alle Situationen eines speziellen Untergraphen in der Liste repräsentiert, so wird mit dem nächst allgemeineren Graphen fortgesetzt, bis alle Situation in der Liste vertreten sind.

Auch die perspektivischen Interpretationen der einzelnen Situationen und Situationsübergänge werden als Liste (Perspektivierungsliste) gespeichert. Jedes Glied repräsentiert dabei genau eine Situation bzw. genau einen Situationsübergang. Die Reihenfolge der Glieder entspricht der Reihenfolge, in der die Situationen bzw. Situationsübergänge in der Datei *perspektive.persp* aufgeführt sind.



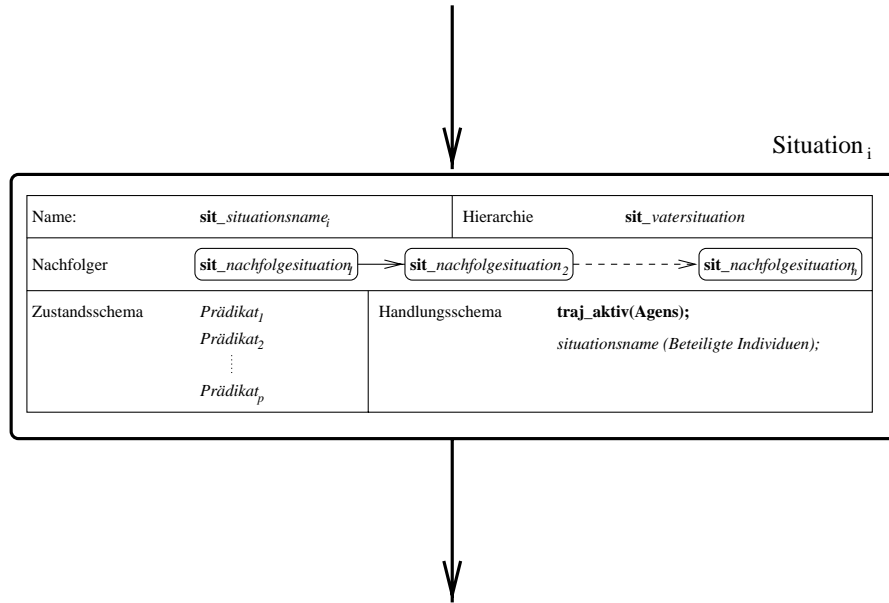


Abbildung C.1: Datenstruktur in MALCOLM.

## C.3 Konstruktion des Perspektivierungsgraphen

### C.3.1 Initialisierung des Perspektivierungsgraphen

Im ersten Schritt wird die Liste, die den Situationsbestimmungsbaum repräsentiert (**Sitgraphliste**), in eine identische Liste kopiert, die nach Abschluß der Berechnungen den Perspektivierungsgraphen (**Perspgraphliste**) enthalten wird. Aus der Perspgraphliste werden dann diejenigen Situationen entfernt, die im Situationsbestimmungsbaum keine Blattsituationen waren, mit Ausnahme der Situationen des Wurzelgraphen. In den verbliebenen Gliedern wird das Feld, das die Nachfolgesituationen der jeweiligen Situation angibt, gelöscht. Die Liste der Situationsnachfolger wird später neu berechnet werden. Ebenso wird das Zustandsschema jeder Situation gelöscht. Die in den Handlungsschemata des Situationsbestimmungsbaums enthaltenen Prädikate werden in die Zustandsschemata des Perspektivierungsbaumes kopiert. Der Hintergrund für diese Maßnahme ergibt sich aus den Ausführungen im Abschnitt 3.4.4.

Im nächsten Schritt wird in der Perspektivierungsliste dasjenige Glied gesucht, welches die perspektivische Beschreibungen der aktuell bearbeiteten Situation des Perspektivierungsbaumes enthält. Die Prädikate der Ebene **1** werden fortan das Handlungsschema der Situation bilden. Zum Hintergrund für diese Maßnahme sei wiederum auf Abschnitt 3.4.4 verwiesen. Enthält die Perspektivierungsliste auch Beschreibungen der Ebene **2**, so wird das aktuelle Glied des Perspgraphliste in ein neues Glied kopiert. Das Zustandsschema beider Glieder ist somit identisch. An den in dem neuen Glied enthaltenen Situationsnamen wird zur Unterscheidung der beiden Glieder die Ziffer '2' angehängt.

Die Prädikate der Ebene **2** bilden fortan das Handlungsschema in dem neuen Glied. Entsprechend wird verfahren, wenn die Perspektivierungsliste auch Beschreibungen der Ebene **3** enthält.

Findet sich in der Perspektivierungsliste eine perspektivische Beschreibung des Übergangs von der im aktuellen Glied repräsentierten Situation auf eine andere Situation (Funktion **finde\_uebergang**), so wird wie folgt verfahren: Hinter das aktuelle Glied wird ein neues Glied eingefügt. Dieses erhält als Situationsnamen den Namen der aktuellen Situation, an den die Zeichenfolge ‘\_to\_’ und anschließend der Name der Situation angehängt wird, zu der übergegangen werden soll. Als Zustandsschema werden in diesem neuen Glied die Prädikate des Handlungsschemas der Situation eingetragen, zu der übergegangen wird. Das Handlungsschema wird aus der perspektivischen Beschreibung aus der Perspektivierungsliste abgeleitet. Zum Hintergrund für diese Maßnahmen sei erneut auf Abschnitt 3.4.4 verwiesen.

### C.3.2 Rekonstruktion der Nachfolgebeziehungen

Nach Abschluß der Initialisierung des Perspektivierungsgraphen bestimmt die Funktion ‘**bilde\_nachfolgekanten**’ die Nachfolgebeziehungen zwischen den einzelnen Situationen des Graphen.

Existiert für eine Situationsbeschreibung des Perspektivierungsbaumes eine Situationsbeschreibung gleichen Namens im Situationsbestimmungsbaum (d.h. enthält erstgenannte Situationsbeschreibung nur Beschreibungen der Ebene ‘1’), so wird die Situationsnachfolge wie folgt geregelt:

- Erster potentieller Nachfolger der Situation ist die Situation selbst.
- Weitere Nachfolger sind diejenigen Situationen, die bereits Nachfolger der korrespondierenden Situation des Situationsbestimmungsbaumes waren, sowie alle deren Kindsituationen.
- Desweiteren werden die potentiellen Nachfolger aller hierarchisch übergeordneten Situationen der korrespondierenden Situation des Situationsbestimmungsbaumes, sowie die Kinder all dieser Nachfolger als potentielle Nachfolger der Situationsbeschreibung im Perspektivierungsbaumes berücksichtigt.

Existiert für eine Situationsbeschreibung des Perspektivierungsbaumes hingegen **keine** Situationsbeschreibung gleichen Namens im Situationsbestimmungsbaum, so sind zwei Ursachen denkbar.

1. Die Situationsbeschreibung enthält Beschreibungen der Ebene ‘2’ oder ‘3’.
2. Die Situationsbeschreibung beschreibt einen Situationsübergang.

Im erstgenannten Fall übernimmt die aktuelle Situationsbeschreibung die Nachfolgebeziehungen derjenigen Situation des Perspektivierungsgraphen, die die korrespondierenden Beschreibungen der Ebene '1' enthält. Als einziger potentieller Nachfolger dieser Situation wird die aktuelle Situation modelliert (siehe Abbildung 3.5 rechts unten).

Im zweiten Fall wird die Situationsnachfolge folgendermaßen geregelt: Einziger potentieller Nachfolger der aktuellen Situation wird die Zielsituation des Situationsübergangs. Einziger Nachfolger der Ausgangssituation wird die aktuelle Situation (siehe Abbildung 3.5 rechts Mitte).

Bei der Durchführung des Algorithmus kann es geschehen, daß ein und dieselbe Situation mehrfach und mit unterschiedlicher Priorität als potentielle Nachfolger einer bestimmten Situation ermittelt wird. Diese Mehrfachbelegungen werden nachträglich gelöscht, indem nur die Nachfolgebeziehung mit der höchsten Priorität bewahrt wird und alle anderen gelöscht werden (Funktion **mehrfachnachfolger\_raus**).

### C.3.3 Rekonstruktion der Hierarchiebeziehungen

Der resultierende Perspektivierungsgraphenbaum besitzt die Tiefe **2**. Der Wurzelgraph des Perspektivierungsgraphenbaums ist mit dem Wurzelgraphen des zugrundeliegenden Situationsbestimmungsbaums identisch. Alle Situationen, die nicht Teil des Wurzelgraphen sind, werden in einem gemeinsamen Graphen zusammengefaßt (Funktion **aendere\_hierarchie**). Dieser Graph wird als Untergraph der Situation **sit\_aktiv** des Wurzelgraphen aufgefaßt. Es wird vorausgesetzt, daß es im Wurzelgraph des Perspektivierungsgraphen (und damit auch im Wurzelgraphen des Situationsbestimmungsgraphen) eine Situation mit diesem Namen gibt.

## C.4 Ausgabe des Perspektivierungsgraphen

Aus der Datenstruktur, die den Perspektivierungsgraphen enthält, wird eine Repräsentation desselben als *SIT++*-Programm abgeleitet. Dieses wird in eine Datei gespeichert. Diese Datei wird im aktuellen Verzeichnis unter dem Namen *perspektive\_baum.sit* angelegt. Sollte bereits eine Datei gleichen Namens bestehen, so wird diese gelöscht (Funktion **schreibe\_sitgraph**).

# Anhang D

## ANGUS

### D.1 Übersicht

Es folgt eine Dokumentation der im Rahmen der vorliegenden Arbeit entwickelten Programmkomponente ANGUS (ausführlicher dokumentiert in [Gerber 00]).

Das Programm ANGUS berechnet aus einer primitiv begrifflichen Beschreibung, die ihrerseits von dem Programmsystem XTRACK ([Kollnig 95], [Haag 98]) berechnet wurde, perspektivische, natürlichsprachliche Szenenbeschreibungen. Der Aufruf von ANGUS umfaßt vier Parameter:

**angus** *bildfolge agenten hauptakteure erzählstil*

Basierend auf diesem Aufruf berechnet ANGUS eine natürlichsprachliche Beschreibung im Erzählstil *erzählstil*  $\in$  {bericht, seb, sep} für die Hauptakteure *hauptakteure* in der Bildfolge *bildfolge*  $\in$  {et, dt, kB, rhein, tank, stau01, stau02, stau03, stau04, stau05, stau06, stau07, stau08, stau09, stau10, stau11, stau12}. In der Beschreibung wird zusätzlich das Verhalten der Agenten *agenten* berücksichtigt, sofern es das Verhalten der Hauptakteure beeinflusst.

Bsp: **angus** et "1 2 3 4 5 6 7 8" "1 2 3 4"seb

berechnet eine natürlichsprachliche Beschreibung des Verhaltens der Fahrzeuge Objekt\_1, Objekt\_2, Objekt\_3 und Objekt\_4 der Ettlinger-Tor-Szene (et) in Form einer szenischen Erzählung aus Beobachtersicht (seb). Bei der Analyse des Verhaltens der Hauptakteure kann außerdem das Verhalten von Objekt\_1 bis Objekt\_8 mit einfließen, sofern es das Verhalten der Hauptakteure beeinflusst.

## D.2 Dateienorganisation

Im folgenden sind die Verzeichnisse angegeben, die ANGUS benötigt, sowie die darin enthaltenen Dateien.

### D.2.1 /home/gerber/src/final

Dieses Verzeichnis enthält die ausführbare Programmdatei ANGUS, ist also dasjenige Verzeichnis, in dem das Programm ANGUS gestartet wird. Desweiteren muß dieses Verzeichnis folgende Dateien enthalten:

Die Dateien **angus.cc**, **angus1.cc**, **angus2.cc**, **angus3.cc**, **angus5.cc** und **header.h** enthalten den Quelltext des Programms ANGUS.

Die **Lexika** sind für die jeweiligen Wortarten enthalten in den Dateien **adjektiv.tex**, **adverb.tex**, **artikel.tex**, **hilfsverb.tex**, **kardinalzahl.tex**, **nomen.tex**, **ordinalzahl.tex**, **personalpronomen.tex**, **praepositionen.tex**, **quantoren.tex** und **verb.tex**. Diese Dateien sind im Abschnitt D.4 näher erläutert.

Die Dateien **grammatik.gra**, **terminalzeichen\_grammatik.tex** und **konstruktionsregeln.kr** enthalten die in der Syntaxanalyse verwendete Grammatik bzw. die zur semantischen Analyse verwendeten Konstruktionsregeln. Näheres dazu in Abschnitt D.5.

Die ausführbare Datei **lim\_angus** enthält folgende beiden Anweisungen:

```
cd /home/gerber/src/limette
limette3 $1 $2 $3 sit sits/superbaum_lim sits/superbaum aspekt anweisungen
```

Die erste Anweisung vollführt einen Wechsel in das angegebene Verzeichnis. Dann wird der Schlußfolgerungsmechanismus LIMETTE mit den angegebenen Parametern aufgerufen. Mehr hierzu entnehme man Abschnitt D.2.3.

Die ausführbare Datei **lim\_angus\_grupp** enthält folgende beiden Anweisungen:

```
cd /home/gerber/src/limette
limette3 sit sits/superbaum_lim perspektive anweisungen_grupp
```

Die erste Anweisung vollführt einen Wechsel in das angegebene Verzeichnis. Dann wird der Schlußfolgerungsmechanismus LIMETTE mit den angegebenen Parametern aufgerufen. Mehr hierzu entnehme man Abschnitt D.2.3. **Lim\_angus\_grupp** wird vom Hauptprogramm von ANGUS aufgerufen.

Die ausführbare Datei **lim\_angus\_persp** enthält folgende beiden Anweisungen:

```
cd /home/gerber/src/limette
limette3 $1 $2 sit sits/superbaum_lim anweisungen
```

Die erste Anweisung vollführt einen Wechsel in das angegebene Verzeichnis. Dann wird der Schlußfolgerungsmechanismus LIMETTE mit den angegebenen Parametern aufge-

rufen. Mehr hierzu entnehme man Abschnitt D.2.3.

## D.2.2 /home/gerber/src/final/bildfolge

Für jede Bildfolge muß im Verzeichnis /**home/gerber/src/final**/ ein Unterverzeichnis *bildfolge* enthalten sein. Derzeit gilt *bildfolge*  $\in$  {et, dt, kB, rhein, tank, stau01, stau02, stau03, stau04, stau05, stau06, stau07, stau08, stau09, stau10, stau11, stau12}. Jedes dieser Verzeichnisse muß folgende Dateien enthalten:

Für jedes verfolgte Fahrzeug **obj<sub>x</sub>** muß eine Datei **lax.lim** existieren, welche die von XTRACK berechneten Limette-Attribute, also die primitiv-begriffliche Beschreibung des Verhaltens des Fahrzeugs **obj<sub>x</sub>**, enthält (siehe [Theilmann 97]).

Außerdem müssen die Dateien **res2.lim** und **rel.lim** enthalten sein, welche die Fahrbahninformation als begriffliche Beschreibung enthalten. Diese Dateien werden ebenfalls von XTRACK im vorhinein berechnet (vgl. [Theilmann 97]). Die Fahrbahnnamen werden in diese Dateien bisher noch interaktiv ergänzt.

Außerdem muß jedes Verzeichnis /**home/gerber/src/final**/*bildfolge* folgende Unterverzeichnisse enthalten:

Im Unterverzeichnis **situationen** speichert ANGUS unter dem Dateinamen **sitx.lim** die Ergebnisse der Situationsanalyse für das Fahrzeug **obj<sub>x</sub>** in LIMETTE-Notation.

Im Unterverzeichnis **perspektive** speichert ANGUS unter dem Dateinamen *erzaehlstil<sub>x</sub>.lim* die perspektivisch begriffliche Beschreibung, die aus dem Situationsanalyseergebnis ableitet wurde.

Im Unterverzeichnis **drsen** speichert ANGUS unter dem Dateinamen **drs\_erzaehlstil<sub>x</sub>.drs.tex** die daraus berechneten DRSEN in LATEX-Notation.

Im Unterverzeichnis **text** speichert ANGUS unter dem Dateinamen *erzaehlstil<sub>x</sub>.txt* die daraus berechneten natürlichsprachlichen Beschreibungen.

## D.2.3 /home/gerber/src/limette

Dieses Verzeichnis enthält die ausführbare Programmdatei LIMETTE ([Schäfer 96]) und zusätzlich die folgenden Dateien:

Die Datei **anweisungen.lim** enthält Anweisungen zur Steuerung der Situationsanalyse (siehe Abschnitt D.6).

Die Datei **anweisungen\_grupp.lim** enthält die Anfragen zur Berechnung erweiterter Schlußfolgerungen aus einer perspektivisch begrifflichen Beschreibung (siehe Abschnitt D.6).

## D.2.4 /home/gerber/src/limette/sits

Dieses Verzeichnis enthält den Situationsgraphenbaum **superbaum.sit** zur Situationsanalyse. Darüber hinaus enthält es zur Berechnung perspektivischer begrifflicher Beschreibungen für jeden Erzählstil  $erzaehlstil \in \{seb, sep, bericht\}$  einen Situationsgraphenbaum  $erzaehlstil.baum.sit$ , die zuvor mit dem Programm MALCOLM berechnet wurden. Die Datei **superbaum.lim.lim** enthält die Definitionen aller LIMETTE-Regeln, die im Situationsgraphen oder bei der Berechnung erweiterter Schlußfolgerungen verwendet werden.

## D.3 Quelltext

Der Quelltext des Programms ANGUS umfaßt folgende Dateien.

```
angus.cc  
angus1.cc  
angus2.cc  
angus3.cc  
angus5.cc  
header.h
```

Die Quelldatei **angus.cc** enthält das Hauptprogramm. Die Datei **angus1.cc** enthält alle Funktionen, die bei der Berechnung einer perspektivischen, begrifflichen Beschreibung aus einer primitiv begrifflichen Beschreibung benötigt werden. Die Datei **angus2.cc** enthält alle Funktionen, die bei der Berechnung einer Diskursrepräsentationsstruktur (DRS, [Kamp&Reyle 93]) aus der perspektivisch begrifflichen Beschreibung benötigt werden. Die Datei **angus3.cc** enthält alle Funktionen, die bei der Ableitung eines natürlichsprachlichen Textes aus der DRS vonnöten sind. Die Datei **angus5.cc** enthält alle Hilfsfunktionen, die von verschiedenen anderen Funktionen aufgerufen werden.

## D.4 Lexika

Zur Syntaxanalyse einer perspektivisch begrifflichen Beschreibung und zur Erzeugung eines natürlichsprachlichen Textes ist lexikalisches Wissen erforderlich. Für jede Wortart existieren in ANGUS eigene Lexika, die teilweise sehr einfach aufgebaut sind. Die Lexika tragen die Dateinamen *wortart.tex*. Ihr Aufbau ist im folgenden dargelegt.

### D.4.1 Verben und Hilfsverben

Ein Verbeintrag in das entsprechende Lexikon umfaßt folgende Merkmale:

1. Die Grundform des Verbes (ohne ‘to’).
2. Das Partizip der Vergangenheit, eingeleitet durch das Schlüsselwort ‘**Past:**’.
3. Die Verlaufsform (‘ing-Form’) des Verbs, eingeleitet durch das Schlüsselwort ‘**Ing:**’.
4. Optional einen Eintrag ‘Bewegungsverb’ oder ‘kein\_Bewegungsverb’, je nachdem, ob es sich um ein Bewegungsverb handelt oder nicht. Dieses Merkmal wird verwendet bei der Assoziation von Temporaladverbien in einen natürlichsprachlichen Text.
5. Eine waagrechte Linie, gebildet aus einer Folge von Minuszeichen, zur Trennung von den Einträgen für das nächste Verb.

## D.4.2 Nomen

Ein Eintrag in das entsprechende Lexikon umfaßt folgende Merkmale:

1. Das Nomen in seiner Grundform.
2. Das Merkmal ‘(Gattung)’ oder ‘(keine Gattung)’ in der darauffolgenden Zeile, je nachdem, ob das betreffende Nomen ein nicht zählbarer oder ein zählbarer Begriff ist. Zum Beispiel ist ‘*petrol*’ im Gegensatz zu ‘*house*’ ein nicht zählbarer, und damit ein Gattungsbegriff. Gattungsbegriffe unterscheiden sich in ihrem sprachlichen Gebrauch u.a. von anderen Nomen dadurch, daß sie oft ohne Artikel geführt werden, wie z.B. in: *The vehicle took petrol*. Deswegen sind sie entsprechend zu kennzeichnen.
3. Eine waagrechte Linie, gebildet aus einer Folge von Minuszeichen, zur Trennung von den Einträgen für das nächste Nomen.

## D.4.3 Kardinalzahlen und Quantoren

Ein Eintrag in das entsprechende Lexikon umfaßt folgende Merkmale:

1. Die Bezeichnung des Quantors.
2. Das Merkmal ‘(Plural)’ oder ‘(Singular)’ in der darauffolgenden Zeile, je nachdem, ob der betreffende Quantor den Plural oder den Singular nach sich zieht.
3. Eine waagrechte Linie, gebildet aus einer Folge von Minuszeichen, zur Trennung von den Einträgen für den nächsten Quantoren.



## D.4.4 Personalpronomen

Die Einträge von Personalpronomen sind hier komplett angegeben:

Fall	Person	Zahl	Geschlecht	Pronomen
NOMINATIV	1	SINGULAR		i
NOMINATIV	2	SINGULAR		you
NOMINATIV	3	SINGULAR	MALE	he
NOMINATIV	3	SINGULAR	FEMALE	she
NOMINATIV	3	SINGULAR	NEUTR	it
NOMINATIV	1	PLURAL		we
NOMINATIV	2	PLURAL		you
NOMINATIV	3	PLURAL		they
AKKUSATIV	1	SINGULAR		me
AKKUSATIV	2	SINGULAR		you
AKKUSATIV	3	SINGULAR	MALE	him
AKKUSATIV	3	SINGULAR	FEMALE	her
AKKUSATIV	3	SINGULAR	NEUTR	it
AKKUSATIV	1	PLURAL		us
AKKUSATIV	2	PLURAL		you
AKKUSATIV	3	PLURAL		them

## D.4.5 Sonstige Wortarten

Die Lexika für *Adjektive*, *Adverbien*, *Artikel*, *Präpositionen* und *Ordinalzahlen* sind besonders einfach aufgebaut. Sie enthalten lediglich die Namen der jeweils zugehörigen Wörter (in alphabetischer Ordnung), paarweise getrennt durch waagrechte Linien, gebildet aus einzelnen Minuszeichen.

## D.5 Konstruktionsregeln und Grammatiken

Die Syntaxanalyse einer perspektivisch begrifflichen Beschreibung wird anhand einer Grammatik vorgenommen. Sie ist enthalten in der Datei **grammatik.gra**, deren Aufbau im folgenden erklärt wird. Bei der Berechnung einer DRS aus Strukturbäumen werden Konstruktionsregeln verwendet, die in der Datei **konstruktionsregeln.kr** enthalten sind. Auch deren Aufbau wird im folgenden erklärt.

## D.5.1 Grammatik

Hier die aktuell verwendete Grammatik (Datei **grammatik.gra**). Konstante Ausdrücke werden durch das Dollarzeichen '\$' eingegrenzt.

```
F => VERBREATION
F => RAUMRELATION
F => QUANTORRELATION
F => METARELATION
VERBREATION => PRAEDNAME $($ AG $,$ BO $,$ MK3 $).$
VERBREATION => PRAEDNAME $($ AG $,$ MK3 $).$
METARELATION => RELNAME $($ MK3 $,$ MK1 $,$ MK2 $).$
METARELATION => RELNAME $($ MK3 $,$ MK1 $).$
RAUMRELATION => RAUMREL $($ MK3 $,$ MK1 $,$ BO $).$
RAUMRELATION => RAUMREL $($ AG $,$ BO $,$ MK3 $).$
QUANTORRELATION => $quantor($ ID $,$ GM $,$ UM $).$
QUANTORRELATION => $quantor($ ID $,$ PRAEDNAME $(X,$ BO $),$ UM $).$
PRAEDNAME => V
PRAEDNAME => V ADV
PRAEDNAME => V PRAEP
PRAEDNAME => AUX V
PRAEDNAME => AUX V PRAEP
PRAEDNAME => V ADJ
PRAEDNAME => V Ving
PRAEDNAME => V Vpast
PRAEDNAME => V ADV ADV
PRAEDNAME => PRAEP Nsing
AG => EN
AG => NPsing
AG => QP
AG => NPplur
AG => PP
NPsing => ADJ Nsing
NPsing => DET Nsing
NPsing => DET ADJ Nsing
NPsing => DET ORD Nsing
NPsing => Nsing
NPplur => ADJ Nplur
NPplur => DET Nplur
QP => CARDsing Nsing
QP => CARDplur Nplur
QP => QUANTsing Nsing
QP => QUANTplur Nplur
BO => EN
```

BO => NPsing  
 BO => MK1  
 BO => NPplur  
 MK1 => MK  
 MK2 => MK  
 MK2 => ADV  
 MK3 => MK  
 MK3 => EN  
 ID => QP  
 GM => PRAEDNAME \$(X)\$  
 GM => PRAEDNAME \$(X,\$ BO \$)\$  
 GM => PRAEDNAME \$(X,Y2)\$  
 UM => PRAEDNAME \$(X,\$ BO \$)\$  
 RAUMREL => PRAEP  
 RELNAME => EN

Die verwendeten Terminalzeichen müssen in der Datei **terminalzeichen\_grammatik.tex** angegeben werden. Dies sind zur Zeit die folgenden:

V	Verb
PRAEP	Praeposition
AUX	Hilfsverb
ADV	Adverb
Ving	Verb in der Verlaufsform
Vpast	Past Partizip eines Verbs
PP	Personalpronomen
ADJ	Adjektiv
DET	Artikel
Nplur	Nomen im Plural
Nsing	Nomen im Singular
CARDplur	Kardinalzahl, die mit dem Plural steht
CARDsing	Kardinalzahl, die den Singular nach sich fuehrt
QUANTplur	Quantor, der mit dem Plural steht
QUANTsing	Quantor, der den Singular nach sich fuehrt
ORD	Ordinalzahl
MK	Markierung
EN	Eigennamen

Es ist zu beachten, daß jede Grammatikregel entweder nur Terminalzeichen oder nur Nichtterminalzeichen enthalten darf.

## D.5.2 Konstruktionsregeln

Die Konstruktionsregeln von ANGUS sind in der Datei **konstruktionsregeln.kr** enthalten. Einträge haben folgenden Aufbau:

- Der erste Eintrag beginnt mit dem Schlüsselwort **'konstruktionsregel'**, gefolgt von dem Namen der entsprechenden Regel, der üblicherweise mit **'KR\_'** beginnt.
- Die auslösende Konfiguration der die Regel wird in einer **'TRIGGER ... END-TRIGGER'** - Umgebung in Form einer Grammatik angegeben.
- Soll die Konstruktionsregel bewirken, daß ein neuer Diskursreferent  $x$  in das Universum der DRS eingefügt werden soll, so wird entsprechendes durch den Eintrag **'NEW\_DR: x'** gewährleistet.
- Mittels des Schlüsselworts **'SET\_CON'** werden neu einzufügende DRS-Bedingungen gekennzeichnet. Beispiel: Soll die Konstruktionsregel bewirken, daß die DRS-Bedingung  $x = y$  in die DRS aufgenommen wird, so geschieht dies durch folgenden Eintrag: **'SET\_CON: x=y'**
- Soll die Eintragung einer neuen DRS-Bedingung in die DRS davon abhängen, ob dort bereits eine bestimmte andere DRS-Bedingung vorhanden ist oder nicht, so wird dies durch das Schlüsselwort **'SEARCH\_IN\_CON'** erreicht. Beispiel:

**SEARCH\_IN\_CON: alpha(DR)**

{

**EX: SET\_CON: x=DR**

**ELSE: SET\_CON: alpha(x)**

}

bewirkt folgendes: Es wird überprüft, ob die DRS eine DRS-Bedingung **'alpha(DR)'** enthält. Ist dies der Fall, so wird die neue DRS-Bedingung **'x=DR'** eingefügt. Andernfalls wird **'alpha(x)'** eingefügt.

- Duplexbedingungen, die durch eine bestimmte Konstruktionsregel erzeugt werden sollen, werden in einer **'DUPLEX ... ENDDUPLEX'** - Umgebung angegeben. Beispiel: Der Eintrag

**DUPLEX**

**4: a**

**1: NEW\_DR: x**

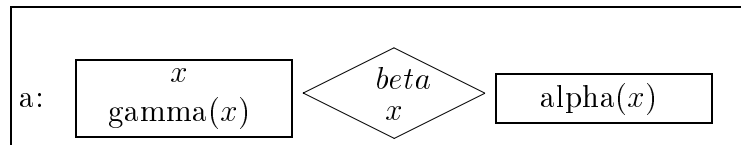
**1: SET\_CON: gamma(x)**

**2: SET\_CON: beta x**

**3: SET\_CON: alpha(x)**

## ENDDUPLEX

erzeugt folgende Duplexbedingung:



Das Schlüsselwort ‘**1:**’ enthält folglich die Eintragungen für die linke innere DRS einer Duplexbedingung, das Schlüsselwort ‘**2:**’ die Eintragungen für die Raute, das Schlüsselwort ‘**3:**’ die Eintragungen für die rechte innere DRS einer Duplexbedingung und das Schlüsselwort ‘**4:**’ den Diskursreferenten, der den in der Duplexbedingung beschriebenen Ablauf repräsentiert. Man beachte, daß die aus dem griechischen Alphabet entlehnten Bezeichnungen *alpha*, *beta* und *gamma* Bezeichner sind, die durch Eintragungen in dem Strukturbaum, auf den die entsprechende Konstruktionsregel angewendet wird, ausgeprägt werden.

- Erzeugt eine Konstruktionsregel eine innere DRS innerhalb der DRS, so wird dies durch eine ‘**INNER\_DRS ... END\_INNER\_DRS**’ - Umgebung angegeben.
- Soll eine Konstruktionsregel bewirken, daß eine bestimmte DRS-Bedingung oder eine innere DRS aus der DRS entfernt wird, so wird dies durch die Schlüsselworte ‘**DELETE\_CON**’ bzw. ‘**DELETE\_INNERDRS**’, gefolgt von der jeweiligen Bedingung bzw. DRS, angezeigt.
- Der Strukturbaum, der nach Anwendung der Konstruktionsregel verbleiben soll, wird in einer ‘**NEW\_TRIGGER ... END\_NEW\_TRIGGER**’ - Umgebung in Form einer Grammatik angegeben.

## D.6 Anbindung von LIMETTE

Das Programm ANGUS ruft an drei verschiedenen Stellen den Schlußfolgerungsmechanismus LIMETTE auf. Dies geschieht zur Situationsanalyse, zur Berechnung einer perspektivisch begrifflichen Beschreibung und zur Berechnung erweiterter Schlußfolgerungen aus letztgenannter Beschreibung.

In allen drei Fällen wird dabei auf die gleiche Weise verfahren. Es wird eine Zeichenkette generiert, die den LIMETTE-Aufruf enthält. Diese Zeichenkette wird durch das “C”-Kommando **system** als Befehl aufgefaßt und ausgeführt.

### D.6.1 limette\_aufruf

In der Funktion **limette\_aufruf** wird der folgende Aufruf gebildet:

```
lim_angus /home/gerber/src/final_angus/bildfolge/res2  
/home/gerber/src/final_angus/bildfolge/rel  
/home/gerber/src/final_angus/bildfolge/daten
```

Die Bedeutung der drei Aufrufparameter, d.h. der Inhalt der drei Dateien wird im Abschnitt D.2 erklärt.

Der Aufruf der ausführbaren Datei **lim\_angus** führt zu einem Wechsel in das Verzeichnis, das den Schlußfolgerungsmechanismus LIMETTE enthält. Dieser wird dann mit folgenden Parametern automatisch aufgerufen:

```
limette3 /home/gerber/src/final_angus/bildfolge/res2  
/home/gerber/src/final_angus/bildfolge/rel  
/home/gerber/src/final_angus/bildfolge/daten  
sit sits/superbaum_lim sits/superbaum aspekt anweisungen
```

Die Bedeutung der meisten dieser Dateien ist in Abschnitt D.2 erklärt. Die Datei **aspekt.lim** wird von ANGUS automatisch erzeugt. Die Datei **anweisungen.lim** enthält die automatisch durchgeführten Anfragen an LIMETTE, steuert also die Situationsanalyse. Sie enthält folgende Anfragen:

```
?- set(defcheck,off). // Ausschalten störender Plausibilitätsprüfungen  
?- tra_faktenberechnung. // Starten der Situationsanalyse  
?- quit. // Verlassen von LIMETTE, Rückkehr zu ANGUS
```

Das Ergebnis der Situationsanalyse wird in die Datei **ergebnis.lim** geschrieben, welche von ANGUS sofort in sein Heimverzeichnis verschoben wird.

## D.6.2 limette\_perspektive

In der Funktion **limette\_perspektive** wird der folgende Aufruf gebildet:

```
lim_angus_persp /home/gerber/src/final_angus/ergebnis  
/home/gerber/src/limette/sits/erzaehlstil_baum
```

Die Bedeutung der Aufrufparameter, d.h. der Inhalt der Dateien wird im Abschnitt D.2 erklärt.

Der Aufruf der ausführbaren Datei **lim\_angus\_persp** führt zu einem Wechsel in das Verzeichnis, das den Schlußfolgerungsmechanismus LIMETTE enthält. Dieser wird dann mit folgenden Parametern automatisch aufgerufen:

```
limette3 /home/gerber/src/final_angus/ergebnis  
/home/gerber/src/limette/sits/erzaehlstil_baum  
sit sits/superbaum_lim anweisungen
```

Die Bedeutung dieser Dateien ist im vorigen Abschnitt oder in Abschnitt D.2 erklärt.

Das Ergebnis der Auswertung wird in die Datei **perspektive.lim** geschrieben.

### **D.6.3 limette\_grupp**

Im Hauptprogramm wird die folgende ausführbare Datei aufgerufen: **lim\_angus\_grupp**. Dieser Aufruf führt zu einem Wechsel in das Verzeichnis, das den Schlußfolgerungsmechanismus LIMETTE enthält. Dieser wird dann mit folgenden Parametern automatisch aufgerufen:

**limette3 sit sits/superbaum\_lim perspektive anweisungen\_grupp**

Die Bedeutung dieser Dateien ist im vorigen Abschnitt oder in Abschnitt D.2 erklärt. Das Ergebnis der Auswertung wird in die Datei **perspektive.lim** geschrieben.

# Literaturverzeichnis

- [Abella&Kender 96] A. Abella, J.R. Kender: *Description Generation of Abnormal Densities Found in Radiographs*. Fourth European Conference on Computer Vision 1996 (ECCV '96), Workshop on Conceptual Descriptions from Images, Cambridge/UK, 19 April 1996, pp. 97-111.
- [Adorni 94] G. Adorni: *Spatial Reasoning as a Tool for Scene Generation and Recognition*. In: V. Cantoni (Ed.), *Human and Machine Vision: Analogies and Divergencies*, Plenum Press, New York, 1994, pp. 289-318.
- [Akman&Surav 96] V. Akman, M. Surav: *Steps toward Formalizing Context*. *AI Magazine* **17**:3 (Fall 1996) 55-72.
- [Allen 83] J.F. Allen: *Maintaining Knowledge about Temporal Intervals*. *Comm. ACM* **26** (1983) 832-843.
- [André et al. 88] E. André, G. Herzog, T. Rist: *The System Soccer*. Proc. 8th Europ. Conference on Artificial Intelligence, Munich / Germany, 1-5 August 1988, pp. 449-454.
- [Bloomberg&Chen 96] D.S. Bloomberg, F.R. Chen: *Document Image Summarization without OCR*. Proc. IEEE International Conference on Image Processing (ICIP '96), Lausanne/CH, 16-19 September 1996, Vol. II, pp. 229-232.
- [Bolle et al. 98] R. M. Bolle, B.-L. Yeo, M. M. Yeung: *Video Query: Research Directions*. *IBM Journal of Research and Development* **42**:2 (March 1998) 233-252.
- [Broszka 94] C. Broszka: *Temporallogisches Programmieren*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), 1994.
- [Burkert 96] G. Burkert: *Repräsentation von lexikalisch-semantischem Wissen in einem System zur Verarbeitung natürlicher Sprache*. Dissertation, Fakultät für Informatik der Universität Stuttgart, Juni 1996. Erschienen in: *Dissertationen zur künstlichen Intelligenz (DISKI)* **141**, infix-Verlag, Sankt Augustin, 1996.
- [Buxton&Gong 95] H. Buxton, S. Gong: *Visual Surveillance in a Dynamic and Uncertain World*. *Artificial Intelligence* **78** (1995) 431-459.



- [Canisius 87] P. Canisius: *Perspektivität in Sprache und Text*. Brockmeyer-Verlag, Bochum, 1987.
- [Charniak 91] E. Charniak: *Bayesian Networks without Tears*. AI Magazine **12**:6 (Winter 1991) 50–63.
- [Chella et al. 97] A. Chella, M. Frixione, S. Gaglio: *A Cognitive Architecture for Artificial Vision*. Artificial Intelligence **89** (1997) 73-111.
- [Clay&Wilhelms 96] S.R. Clay, J. Wilhelms: *Put: Language Based Interactive Manipulation of Objects*. IEEE Computer Graphics and Applications **16**:2 (March 1996) 31-39.
- [Courtney 97] J. D. Courtney: *Automatic Video Indexing via Object Motion Analysis*. Pattern Recognition **30**:4 (1997) 607-625.
- [Damm 93] H. Damm: *Zur Übertragung eines Systems zur Extraktion von Bewegungsverben aus Bildfolgen in die englische Sprache*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, November 1993.
- [Damm 94] H. Damm: *Entwurf und Implementierung von Situations- und Handlungsschemata zur Beschreibung von Abläufen einer Tankstellenszene*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Dezember 1994.
- [Dance et al. 96] S. Dance, T. Caelli, Z.-Q. Liu: *A Concurrent, Hierarchical Approach to Symbolic Scene Interpretation*. Pattern Recognition **29**:11 (1996) 1891–1903.
- [Eberle 1997] K. Eberle: *Flat Underspecified Representation and its Meaning for a Fragment of German*. Sprachtheoretische Grundlagen für die Computerlinguistik, Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 120, Oktober 1997.
- [van Eijck&Kamp 97] J. van Eijck, H. Kamp: *Representing Discourse in Context*. J. van Benthem, A. ter Meulen (Eds.): Handbook of Logic and Language, Elsevier Science B.V. Amsterdam a.o., 1997, 179-237.
- [Fautz 96] M. Fautz: *Automatisiertes Anpassen eines Kreuzungsmodells an Teile von innerstädtischen Kreuzungen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), März 1996.
- [Fernyhough et al. 98] J. Fernyhough, A. G. Cohn, D. C. Hogg: *Event Recognition using Qualitative Reasoning on Automatically Generated Spatio-Temporal Models from Visual Input*. H. Buxton and A. Mukerjee (Eds.): Proc. ICCV-98 Workshop on Conceptual Descriptions of Images (CDI-98), Bombay/India, 2 January 1998, pp. 39-47.

- [Frank 95] T. Frank: *Erkennung und Verfolgung von verdeckten Fahrzeugen in Straßenverkehrsszenen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), August 1995.
- [Frank et al. 96] T. Frank, M. Haag, H. Kollnig, H.-H. Nagel: *Tracking of Occluded Vehicles in Traffic Scenes*. Fourth European Conference on Computer Vision 1996 (ECCV'96), Vol. II, Cambridge, England, 15-18 April 1996, Lecture Notes in Computer Science **1065**, Springer-Verlag, Berlin a.o. 1996, pp. 485-494.
- [Friedman 96] L. Friedman: *From Images to Language*. Proc. Workshop on Conceptual Descriptions from Images, Cambridge/UK, 19 April 1996, pp. 70-81.
- [Früauf 94] T. Früauf: *Berechnung begrifflicher Beschreibungen für Ansammlungen von Fahrzeugen durch Bildfolgenauswertung*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Dezember 1994.
- [Gapp 94b] K.-P. Gapp: *Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space*. Bericht Nr. 101, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, April 1994. Erschienen in: Proc. 12th National Conference on Artificial Intelligence (AAAI '94), Seattle, Washington, USA, 2-3 August 1994, pp. 1393-1398.
- [Gapp&Maaß 94] K.-P. Gapp, W. Maaß: *Spatial Layout Identification and Incremental Descriptions*. Bericht Nr. 102, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Mai 1994. Erschienen in: P. Mc Kevitt (Ed.), Proc. Workshop on the Integration of Natural Language and Vision Processing, 12th National Conference on Artificial Intelligence (AAAI '94), Seattle, Washington, USA, 2-3 August 1994, pp. 145-152.
- [Gerber 94] R. Gerber: *Diskurs-Repräsentationstheorie zur Darstellung begrifflicher Beschreibungen von Abläufen in Straßenverkehrsszenen*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Dezember 1994.
- [Gerber 95] R. Gerber: *Diskurs-Repräsentationstheorie zur Situationsanalyse in Straßenverkehrsszenen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Juli 1995.
- [Gerber&Nagel 96a] R. Gerber, H.-H. Nagel: *Berechnung natürlichsprachlicher Beschreibungen von Straßenverkehrsszenen aus Bildfolgen unter Verwendung von Geschehens- und Verdeckungsmodellierung*. In B. Jähne, P. Geißler, H. Haußecker und F. Hering (Hrsg.), Mustererkennung 1996, 18. DAGM-Symposium, Heidelberg, 11.-13. September 1996, Springer-Verlag, Berlin, Heidelberg u.a., pp. 601-608.

- [Gerber&Nagel 96b] R. Gerber, H.-H. Nagel: *Knowledge Representation for the Generation of Quantified Natural Language Descriptions of Vehicle Traffic in Image Sequences*. Proc. IEEE International Conference on Image Processing (ICIP '96), Special Session on Knowledge-Intensive Image Interpretation, Lausanne/CH, 16-19 September 1996, Vol. II, pp. 805-808.
- [Gerber&Nagel 98] R. Gerber, H.-H. Nagel: *(Mis?)-Using DRT for Generation of Natural Language Text from Image Sequences*. Proc. European Conference on Computer Vision (ECCV '98), Freiburg/Germany, 2-6 June 1998, Vol. II, pp. 255-270.
- [Gerber 99] R. Gerber: *Einige Überlegungen zu den Anmerkungen von Herrn Professor Kamp vom 12.05.1999 bezüglich unserer Arbeit*. Interner Bericht, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Juni 1999.
- [Gerber 00] R. Gerber: *ANGUS Dokumentation*. Interner Bericht, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Februar 2000.
- [Grewendorf et al. 91] G. Grewendorf, F. Hamm, W. Sternefeld: *Sprachliches Wissen: Eine Einführung in moderne Theorien der grammatischen Beschreibung*. Suhrkamp-Verlag, Frankfurt/Main, 1991.
- [Guthrie et al. 96] L. Guthrie, J. Pustejovsky, Y. Wilks, B.M. Sator: *The Role of Lexicons in Natural Language Processing*. Communications of the ACM **39**:1 (January 1996) 63-72.
- [Haag 94] M. Haag: *Modellierung eines Situationsgraphen am Beispiel einer Tankstellenszene*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Mai 1994.
- [Haag 95] M. Haag: *Situationsgraphen zur Modellierung von Zuständen und Handlungen im Diskursbereich Straßenverkehr*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), März 1995.
- [Haag 98] M. Haag: *Bildfolgenauswertung zur Erkennung der Absichten von Straßenverkehrsteilnehmern*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Juli 1998. Erschienen in: Dissertationen zur künstlichen Intelligenz (DISKI) **193**, infix-Verlag, Sankt Augustin, Juli 1998.
- [Haag&Nagel 97] M. Haag, H.-H. Nagel: *Verfahren zur modellgestützten Verfolgung und Beschreibung von Objektbewegungen in Bildfolgen*. Interner Bericht, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), März 1997.

- [Haag&Nagel 98a] M. Haag, H.-H. Nagel: *Tracking of Complex Driving Manoeuvres in Traffic Image Sequences*. *Image and Vision Computing* **16**:8 (1998) 517-527.
- [Haag&Nagel 98b] M. Haag und H.-H. Nagel: *'Begriffliche Rückkopplung' zur Behandlung temporärer Verdeckungssituationen in der Bildfolgenauswertung von Straßenverkehrsszenen*. 28. Jahrestagung der Gesellschaft für Informatik 1998 (Informatik'98), 21.-25. September 1998, Magdeburg, J. Dassow und R. Kruse (Hrsg.) *Informatik aktuell*, Springer-Verlag Berlin, Heidelberg 1998, pp.13-22.
- [Haag&Nagel 98c] M. Haag, H.-H. Nagel: *Incremental Recognition of Traffic Situations from Video Image Sequences*. H. Buxton and A. Mukerjee (Eds.): *Proc. ICCV-98 Workshop on Conceptual Descriptions of Images (CDI-98)*, Bombay/India, 2 January 1998, pp. 1-20.
- [Haag&Nagel 99] M. Haag, H.-H. Nagel: *Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Image Sequences*. *International Journal of Computer Vision* **35**:3 (1999) 295-319.
- [Haag et al. 97] M. Haag, T. Frank, H. Kollnig, H.-H. Nagel: *Influence of an Explicitly Modelled 3D Scene on the Tracking of Partially Occluded Vehicles*. *Computer Vision and Image Understanding* **65**:2 (1997) 206-225.
- [Harris 85] M.D. Harris: *Introduction to Natural Language Processing*. Reston Publishing Company, Reston, Virginia/USA 1985.
- [Hearst 97] M. A. Hearst: *Banter on Bayes: Debating the Usefulness of Bayesian Approaches to Solving Practical Problems*. *IEEE Intelligent Systems and their Applications* **12**:6 (November/December 1997) 18-25.
- [Heinze et al. 91] N. Heinze, W. Krüger, H.-H. Nagel: *Berechnung von Bewegungsverbänden zur Beschreibung von aus Bildfolgen gewonnenen Fahrzeugtrajektorien in Straßenverkehrsszenen*. *Informatik Forschung und Entwicklung* **6** (1991) 51-61.
- [Herzog 92] G. Herzog: *Visualization Methods for the VITRA Workbench*. Memo Nr. 53, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Dezember 1992.
- [Herzog&Rohr 95] G. Herzog, K. Rohr: *Integrating Vision and Language: Towards Automatic Description of Human Movements*. Bericht Nr. 122, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Juli 1995. Erschienen in: *Proc. 19th Annual German Conference on Artificial Intelligence*, Bielefeld, 11-13 September 1995, *Lecture Notes in Artificial Intelligence* **981**, Springer-Verlag, Berlin a.o. 1995, pp. 257-268.
- [Herzog&Wazinski 94] G. Herzog, P. Wazinski: *Visual TRANslator: Linking Perceptions and Natural Language Descriptions*. Bericht Nr. 100, KI-Labor am Lehrstuhl

- für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, April 1994. Erschienen in: *Artificial Intelligence Review Journal* **8:2** (1994) 175-187, Special Volume on the Integration of Natural Language and Vision Processing.
- [Howarth&Buxton 98] R. J. Howarth, H. Buxton: *Conceptual Descriptions from Monitoring and Watching Image Sequences*. H. Buxton and A. Mukerjee (Eds.): Proc. ICCV-98 Workshop on Conceptual Descriptions of Images (CDI-98), Bombay/India, 2 January 1998, pp. 21-38.
- [Huang et al. 94] T. Huang, D. Koller, J. Malik, G. Ogasawara, B. Rao, S. Russell, and J. Weber: *Automatic Symbolic Traffic Scene Analysis Using Belief Networks*. Proc. 12th National Conference on Artificial Intelligence, Seattle/WA, 31 July – 4 August 1994, pp. 966–972.
- [Intille 98] S. S. Intille: *Visual Recognition of Multi-agent Action*. Internal Report, Media Lab, Massachusetts Institute of Technology, Cambridge/MA, September 1998.
- [Jameson et al. 94] A. Jameson, B. Kipper, A. Ndiaye, R. Schäfer, J. Simons, T. Weis, D. Zimmermann: *Cooperating to be Noncooperative: The Dialog System PRAC-MA*. Bericht Nr. 106, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Juni 1994. Erschienen in: B. Nebel, L. Dreschler-Fischer (Eds.): KI-94: Advances in Artificial Intelligence, Proc. 18th German Conference on Artificial Intelligence, Saarbrücken, 18-23 September 1994, pp. 106-117.
- [Jeyakumar 98] V. Jeyakumar: *Generation of Synthetic Image Sequences for Car Maneouvers Extracted by Image Sequence Evaluation*. Master Thesis, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Dezember 1998.
- [Jones&Galliers 96] K. S. Jones, J. R. Galliers: *Evaluating Natural Language Processing Systems: An Analysis and Review*. Lecture Notes in Artificial Intelligence **1083**, Springer-Verlag, Berlin a.o. 1996.
- [Kamp&Reyle 93] H. Kamp, U. Reyle: *From Discourse to Logic*. Kluwer Academic Publishers, Dordrecht, Boston/MA, London 1993.
- [Kamp&Rossdeutscher 94a] H. Kamp, A. Rossdeutscher: *Remarks on Lexical Structure and DRS Construction*. *Theoretical Linguistics* **20:2/3** (1994) 98-164.
- [Kamp&Rossdeutscher 94b] H. Kamp, A. Rossdeutscher: *DRS-Construction and Lexically Driven Inference*. *Theoretical Linguistics* **20:2/3** (1994) 165-235.
- [Kamp&Schiehlen 99] H. Kamp, M. Schiehlen: *Vorschlag für die Verbesserung natürlichsprachlicher Beschreibungen, die aus durch Bildverarbeitung gewonnenen Repräsentationen von Verkehrsszenen gewonnen werden*. Interner Bericht, Institut für Maschinelle Sprachverarbeitung, Fakultät für Philosophie der Universität Stuttgart, Mai 1999.

- [King 96] M. King: *Evaluating Natural Language Processing Systems*. Communications of the ACM **39**:1 (January 1996) 73-79.
- [Koller et al. 93] D. Koller, K. Daniilidis, H.-H. Nagel: *Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes*. International Journal of Computer Vision **10** (1993) 257-281.
- [Kollnig 95] H. Kollnig: *Ermittlung von Verkehrsgeschehen durch Bildfolgenauswertung*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Februar 1995. Erschienen in der Reihe: Dissertationen zur Künstlichen Intelligenz (DISKI) **88**, infix-Verlag, Sankt Augustin, 1995.
- [Kollnig&Nagel 93] H. Kollnig, H.-H. Nagel: *Ermittlung von begrifflichen Beschreibungen von Geschehen in Straßenverkehrsszenen mit Hilfe unscharfer Mengen*. Informatik Forschung Entwicklung **8** (1993) 186-196.
- [Kollnig&Nagel 97] H. Kollnig, H.-H. Nagel: *3D Pose Estimation by Directly Matching Polyhedral Models to Gray Value Gradients*. International Journal of Computer Vision **23**:3 (1997) 283-302.
- [Kollnig et al. 95] H. Kollnig, H. Damm, H.-H. Nagel, M. Haag: *Zuordnung natürlicher sprachlicher Begriffe zu Geschehen an einer Tankstelle*. G. Sagerer, S. Posch, F. Kummert (Hrsg.): Mustererkennung 1995, 17. DAGM-Symposium, Bielefeld, 13.-15. September 1995, Springer-Verlag, Berlin u.a., 1995, 236-243.
- [Krüger 91] W. Krüger: *Begriffsgraphen zur Situationsmodellierung in der Bildfolgenauswertung*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Januar 1991. Erschienen in: Informatik-Fachberichte **311**, Springer-Verlag, Berlin u.a., 1992.
- [Leong 98] T. Y. Leong: *Multiple Perspective Dynamic Decision Making*. Artificial Intelligence **105** (1998) 209-261.
- [Lindemann 87] B. Lindemann: *Einige Fragen an eine Theorie der sprachlichen Perspektivierung*. In: P. Canisius (Hrsg.): *Perspektivität in Sprache und Text*. Brockmeyer-Verlag, Bochum, 1987, 1-51.
- [Lüth et al. 94] T. Lüth, T. Längle, G. Herzog, E. Stopp, U. Rembold: *KANTRA: Human-Machine Interaction for Intelligent Robots Using Natural Language*. Bericht Nr. 104, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Juni 1994. Erschienen in: Proc. 3rd International Workshop on Robot and Human Communication (RO-MAN '94), Nagoya/Japan, 18-20 July 1994, pp. 106-111.
- [Maaß 94] W. Maaß: *From Vision to Multimodal Communication: Incremental Route Descriptions*. Bericht Nr. 108, KI-Labor am Lehrstuhl für Informatik IV (Prof.

- Wahlster), Saarbrücken, SFB 314, Juli 1994. Erschienen in: *Artificial Intelligence Review Journal* **8:2** (1994) 159-174, Special Volume on the Integration of Natural Language and Vision Processing.
- [Mück 94] K. Mück: *Entwurf und Anpassung begrifflicher und geometrischer Modelle an Bildfolgen von innerstädtischen Straßenkreuzungen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), November 1994.
- [Nagel 77] H.-H. Nagel: *Analyzing Sequences of TV Frames*. Proc. 4th International Joint Conference on Artificial Intelligence IJCAI '77, Cambridge/MA, 22-25 August 1977, p. 626.
- [Nagel 79] H.-H. Nagel: *Über die Repräsentation von Wissen zur Auswertung von Bildern*. Informatik-Fachberichte **20**, 1. DAGM-Symposium, Karlsruhe, 10.-12. Oktober 1979, pp. 3-21.
- [Nagel 81] H.-H. Nagel: *Image Sequence Analysis: What Can We Learn From Applications?* T. Huang (Ed.): *Image Sequence Analysis*, Springer-Verlag, Berlin u.a., 1981, pp. 19-228.
- [Nagel 88] H.-H. Nagel: *From Image Sequences towards Conceptual Descriptions*. *Image and Vision Computing* **6:2** (1988) 59-74.
- [Nagel 91] H.-H. Nagel: *La représentation de situations et leur reconnaissance à partir de séquences d'images - The Representation of Situations and their Recognition from Image Sequences*. 8. Congrès Reconnaissance des Formes et Intelligence Artificielle, Lyon-Villeurbanne, 25-29 Novembre 1991, AFCET, 1991, pp. 1221-1229.
- [Nagel 99] H.-H. Nagel: *Natural Language Description of Image Sequences as a Form of Knowledge Representation*. W. Burgard, T. Christaller, A. Cremers (Eds.): *KI-99: Advances in Artificial Intelligence*, Proc. 23rd Annual German Conference on Artificial Intelligence, Lecture Notes in Artificial Intelligence **1701**, Bonn, 13-15 September 1999, pp. 45-60.
- [Nagel et al. 95] H.-H. Nagel, H. Kollnig, M. Haag, H. Damm: *The Association of Situation Graphs with Temporal Variations in Image Sequences*. Working Notes AAAI-95 Fall Symposium Series 'Computational Models for Integrating Language and Vision', Cambridge/MA, 10-12 November 1995, pp. 1-8.
- [Nagel et al. 99] H.-H. Nagel, M. Haag, V. Jeyakumar, A. Mukerjee: *Visualisation of Conceptual Descriptions Derived from Image Sequences*. Mustererkennung 1999, 21. DAGM-Symposium, Bonn, 15.-17. September 1999, Springer-Verlag, Berlin, Heidelberg u.a., pp. 364-371.
- [Ndiaye 96] A. Ndiaye: *Rollenübernahme in einem Dialogsystem*. *Künstliche Intelligenz* **4** (1996) 34-40.

- [Neumann&Novak 86] B. Neumann, H.-J. Novak: *NAOS: Ein System zur natürlicher-sprachlichen Beschreibung zeitveränderlicher Szenen*. Informatik Forsch. Entw. **1** (1986) 83-92.
- [Noy&Hafner 97] N.F. Noy, C.D. Hafner: *The State of the Art in Ontology Design*. AI Magazine **18**:3 (1997) 53-74.
- [Pearl 88] J. Pearl: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo/CA, 1988.
- [Pernozzoli 96] A. Pernozzoli: *Diskursrepräsentationstheorie zur Beschreibung von Verdeckungssituationen in Straßenverkehrsszenen*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Februar 1996.
- [Quasthoff 87] U. M. Quasthoff: *Dabeisein durch Sprache: Zur Rolle der Perspektive beim konversationellen Erzählen*. In P. Canisius (Hrsg.): *Perspektivität in Sprache und Text*. Brockmeyer-Verlag, Bochum, 1987, 129-151.
- [Retz-Schmidt 92] G. Retz-Schmidt: *Die Interpretation des Verhaltens mehrerer Akteure in Szenenfolgen*. Dissertation, Universität Saarbrücken. Erschienen in: Informatik-Fachberichte **308**, Springer-Verlag, Berlin u.a., 1992.
- [Reyle 95] U. Reyle: *Co-Indexing Labeled DRSs to Represent and Reason with Ambiguities*. K. van Deemter, S. Peters (Eds.): *Semantic Ambiguity and Underspecification*, CSLI Publications, Stanford/CA, 1995, 239-268.
- [Satoh et al. 97] S. Satoh, Y. Nakamura, and T. Kanade: *Name-It: Naming and Detecting Faces in Video by the Integration of Image and Natural Language Processing*. Proc. 15th International Joint Conference on Artificial Intelligence (IJCAI '97), 23-29 August 1997, Nagoya/Japan, Vol. II, pp. 1488-1493.
- [Schäfer 93] K. Schäfer: *Entwicklung einer temporallogischen Sprache zur Beschreibung von Abläufen in Straßenverkehrsszenen*. Diplomarbeit, Institut für Logik, Komplexität und Deduktionssysteme sowie Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Januar 1993.
- [Schäfer 96] K. Schäfer: *Unschärfe zeitlogische Modellierung von Situationen und Handlungen in Bildfolgenauswertung und Robotik*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Juli 1996. Erschienen in: *Dissertationen zur künstlichen Intelligenz (DISKI)* **135**, infix-Verlag, Sankt Augustin, August 1996.
- [Schirra 94] J. Schirra: *Bildbeschreibung als Verbindung von visuellem und sprachlichem Raum*. Dissertation, April 1994. Erschienen in: *Dissertationen zur Künstlichen Intelligenz (DISKI)* **71**, infix-Verlag, Sankt Augustin, 1994.



- [Schirra&Stopp 93] J. Schirra, E. Stopp: *ANTLIMA - A Listener Model with Mental Images*. In R. Bajcsy (Ed.): Proc. of the 13th International Joint Conference on Artificial Intelligence, Chambéry/France, 28 August - 3 September 1993, pp. 175-180.
- [Schwarz 97] T. Schwarz: *Modellierung und modellgestützte Verfolgung mehrgliedriger Objekte in Videobildfolgen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), August 1997.
- [Smith&Kanade 97] M.A. Smith and T. Kanade: *Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques*. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '97), 17-19 June 1997, Puerto Rico, pp. 775-781.
- [Socher et al. 98] G. Socher, G. Sagerer, P. Perona: *Bayesian Reasoning on Qualitative Descriptions from Images and Speech*. H. Buxton and A. Mukerjee (Eds.): Proc. ICCV-98 Workshop on Conceptual Descriptions of Images (CDI-98), Bombay/India, 2 January 1998, pp. 63-82.
- [Sowa 84] J.F. Sowa: *Conceptual Structures - Information Processing in Mind and Machine*. Addison-Wesley Publishing Company, Reading/MA 1984.
- [Srihari 95] R. K. Srihari: *Linguistic Context in Vision*. Proc. Workshop on Context-Based Vision, Cambridge/MA, 19 June 1995, pp. 100-110.
- [Stockmar 99] G. Stockmar: *Untersuchung von Parallelverarbeitung bei der Bildfolgenauswertung von Verkehrsszenen*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Juni 1999.
- [Stopp 92] E. Stopp: *GEO-ANTLIMA: Konstruktion dreidimensionaler mentaler Bilder aus sprachlichen Szenenbeschreibungen*. Memo Nr. 60, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Universität Saarbrücken, SFB 314, Oktober 1993.
- [Stopp & Blocher 97] E. Stopp, A. Blocher: *Construction of Mental Images and their Use in a Listener Model*. Bericht Nr. 137, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Universität Saarbrücken, Januar 1997.
- [Stopp et al. 94] E. Stopp, K.-P. Gapp, G. Herzog, T. Längle, T. Lüth: *Utilizing Spatial Relations for Natural Language Access to an Autonomous Mobile Robot*. Bericht Nr. 107, KI-Labor am Lehrstuhl für Informatik IV (Prof. Wahlster), Saarbrücken, SFB 314, Juli 1994. Erschienen in: B. Nebel, L. Dreschler-Fischer (Eds.), Proc. 18th German Conference on Artificial Intelligence, Saarbrücken, 18-23 September 1994, pp. 39-50.

[Theilmann 97] W. Theilmann: *Bildfolgenbasierte Beschreibung von Fahrzeugmanövern mit Hilfe unscharfer metrisch temporaler Logik*. Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), März 1997.

[Zeier 99] R. Zeier: *Modellierung von Situationsfolgen*. Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Juni 1999.