Von visuellen Daten zu inkrementellen Wegbeschreibungen in dreidimensionalen Umgebungen: Das Modell eines kognitiven Agenten

Dissertation
zur Erlangung des Grades
des Doktors der Ingenieurswissenschaften
der Technischen Fakultät
der Universität des Saarlandes

von

Wolfgang Maaß

Saarbrücken
1996
Inhaltsverzeichnis

1 Einführung................................................................................. 1
  1.1 Problemstellung............................................................... 2
  1.2 Einordnung der Arbeit...................................................... 2
  1.3 Zielsetzung und Methodologie.......................................... 5
  1.4 Gliederung....................................................................... 7

2 Theorien zur Raumkognition......................................................... 9
  2.1 Zusammenhang zwischen Umgebungen und Situationen........ 19
  2.2 Räumliches Wissen in Situationen.................................... 20
  2.3 Raumkonzepte.................................................................... 21
    2.3.1 Räumliche Referenzsysteme....................................... 24
    2.3.2 Räumliche Relationen................................................ 31
    2.3.3 Wahl eines Referenzsystems im Kontext von Raumbeschreibungen.................... 34
    2.3.4 Landmarkenwissen.................................................... 35
    2.3.5 Routenwissen............................................................ 37
    2.3.6 Konfigurationswissen................................................. 38
    2.3.7 Theorien zu Kognitiven Karten.................................... 38
  2.4 Aspekte des motorischen Verhaltens................................... 39
  2.5 Interaktion von Sprache und räumlichem Wissens............... 40
    2.5.1 Raumdeixis............................................................... 43
    2.5.2 Diskursbereich der Wegbeschreibungen...................... 43
      2.5.2.1 Wegbeschreibungen als Diskurstyp....................... 44
      2.5.2.2 Linearisierung durch reale und imaginierte Wanderungen........ 45
      2.5.2.3 Referentielle Ausdrücke in Wegbeschreibungen............ 47
  2.6 Modelle allgemeiner raumkognitiver Leistungen.................... 48
    2.6.1 Das TOUR Modell.................................................. 48
    2.6.2 Das Modell ELMER.................................................. 49
    2.6.3 Das Modell von Elliott und Lesk............................... 50
    2.6.4 Das Modell TRAVELLER............................................ 50
    2.6.5 Das Modell NAVIGATOR........................................... 50
    2.6.6 Das Basic Agent-Modell.......................................... 52
    2.6.7 Das Modell SPAM.................................................... 53
    2.6.8 Das Modell MERCATOR............................................ 53
INHALTSVERZEICHNIS

2.6.9 Das Modell von Yeap ........................................ 54
2.6.10 Zusammenfassung raumverarbeitender Modelle .......... 54
2.7 Modelle zur Generierung von Wegbeschreibungen .......... 57
  2.7.1 Das formale Modell von Habel ............................. 57
  2.7.2 Das formale Modell von Höök und Karlgren ............... 58
  2.7.3 Das formale Modell von Hoeppner et al. .................... 59
  2.7.4 Das formale Modell von Couckels ......................... 60
  2.7.5 Das Modell von Rau und Schweitzer ....................... 61
  2.7.6 Das Berechnungsmodell KLEIST ............................ 62
  2.7.7 Das Berechnungsmodell CITYGUIDE ......................... 63
  2.7.8 Das Berechnungsmodell von Carstensen .................... 64
  2.7.9 Das Back Seat Driver Modell .............................. 66
  2.7.10 Zusammenfassung der Wegbeschreibungsmoodelle .......... 66
2.8 Zusammenfassung .............................................. 68

3 Empirische Untersuchungen ........................................ 71
  3.1 Vorbemerkungen ............................................. 71
  3.2 Explorative Untersuchung in realen Umgebungen ............ 74
    3.2.1 Arbeitsmodell ......................................... 74
    3.2.2 Durchführung .......................................... 75
    3.2.3 Ergebnis ................................................ 75
    3.2.4 Diskussion ............................................. 81
  3.3 Untersuchung in einer computeranimierten 3D-Umgebung .... 82
    3.3.1 Arbeitsmodell .............................................
    3.3.2 Durchführung .............................................
      3.3.2.1 Methode: Farbvariation ......................... 85
      3.3.2.2 Ergebnis: Farbvariation ......................... 85
      3.3.2.3 Methode: Höhenvariation ..................... 87
      3.3.2.4 Ergebnis: Höhenvariation ..................... 87
    3.3.3 Gesamtdiskussion ....................................... 88

4 Formale Beschreibung des Agenten ................................ 91
  4.1 Modularer Aufbau des funktionalen Modells .................... 92
    4.1.1 Das Steuerungsmodul .................................... 93
    4.1.2 Das Pfadsuchmodul ..................................... 94
    4.1.3 Das Objektauswahlmodul ................................ 96
    4.1.4 Das Raumintegrationsmodul ................................
      4.1.4.1 Räumliche Referenzsysteme ....................... 99
      4.1.4.2 Evaluation räumlicher Relationen ............. 100
      4.1.4.3 Räumlicher Kurzzeitspeicher ................... 102
    4.1.5 Das Navigationsmodul .................................. 102
    4.1.6 Das Sprachproduktionsmodul ............................ 103
4.2 Funktionale Beschreibung ........................................ 103
  4.2.1 Funktionale Beschreibung des Steuerungsmoduls .......... 105
    4.2.1.1 Zeitrahmen und Zeitbeschränkungen .................. 106
    4.2.1.2 Dekomposition des Zeitrahmens \( T_{\text{proc}} \) .......... 108
    4.2.1.3 Aktivation und Deaktivierung von Phasen .......... 110
    4.2.1.4 Aktivierung und Evaluation von Ablaufplänen .......... 112
    4.2.1.5 Funktionen des Steuerungsmoduls .................. 115
  4.2.2 Funktionale Beschreibung des Pfadsuchmoduls ............... 117
    4.2.2.1 Funktionen der inkrementellen heuristischen Pfadsuche ... 118
    4.2.2.2 Pfadabschnitte und Verzweigungspunkte ........... 121
    4.2.2.3 Expansion von Verzweigungspunkten ................ 122
    4.2.2.4 Zusammenfassung ................................ 124
  4.2.3 Funktionen des Objektauswahlmoduls ........................ 125
    4.2.3.1 Allgemeine Beschreibung der Auswahl visuell salienter Objekte .. 125
    4.2.3.2 Identifikation sichtbarer Objekte .................. 127
    4.2.3.3 Abbildung der perspektivischen Sicht auf die Projektionsebene .. 127
    4.2.3.4 Berücksichtigung der räumlichen Fokussierung .......... 129
    4.2.3.5 Abbildung der Projektions- auf die Lokationsebene .......... 130
    4.2.3.6 Bestimmung visueller Salienzen bezüglich des rekursiven Merk-
      mals „Farbe“ ........................................ 132
    4.2.3.7 Bestimmung visueller Salienzen der linearen Merkmale „Höhe“ und „Breite“ .................. 137
    4.2.3.8 Integration in globale Merkmalskarten ................ 138
    4.2.3.9 Adaptation der visuellen Salienzen an den räumlichen Fokusbereich 140
    4.2.3.10 Zusammenfassung ................................ 141
  4.2.4 Funktionen des Raumintegrationsmoduls ..................... 141
    4.2.4.1 Egozentrische und allozentrische räumliche Referenzsysteme .......... 141
    4.2.4.2 Evaluation räumlicher Relationen .................. 150
    4.2.4.3 Der räumliche Kurzzeitsspeicher .................. 155
  4.2.5 Funktionen des Navigationsmoduls .......................... 166
  4.2.6 Funktionen des Sprachproduktionsmoduls .................... 168
    4.2.6.1 Generische Inhaltsstruktur inkrementeller Wegbeschreibungen .. 169
    4.2.6.2 Auswahl von Inhaltsstrukturen bezüglich einer Situation .......... 170
    4.2.6.3 Beispiele der Verwendung räumlicher Konfigurationsbeschrei-
      bungen zur Generierung inkrementeller Wegbeschreibungen . 182
    4.2.6.4 Der Präsentationsspeicher ........................ 182
  4.3 Implementation ............................................. 185
    4.3.1 Prozesse der inkrementellen Pfadsuche .................. 186
    4.3.2 Prozesse zur visuellen Objektauswahl .................. 190
    4.3.3 Prozesse zur Verarbeitung räumlichen Wissens ........... 194
    4.3.4 Prozesse zur Generierung natürlicher Sprache ........... 195
    4.3.5 Beispiel einer automatisch generierten inkrementellen Wegbeschreibung .... 196
5 Ergebnisse und Ausblick 199
  5.1 Wissenschaftlicher Beitrag dieser Arbeit .............................. 199
  5.2 Einordnung in den technologischen Kontext .......................... 201
    5.2.1 Navigationshilfesystem in Realumgebungen ...................... 203
    5.2.2 Navigationshilfe in virtuellen Umgebungen ..................... 208
  5.3 Erweiterungsmöglichkeiten ............................................. 210
### Abbildungsverzeichnis

<p>| 2.1 | feature-integration-theory | 12 |
| 2.2 | ebenen | 14 |
| 2.3 | gollodge-concepts | 22 |
| 2.4 | is-a-concepts | 23 |
| 2.5 | frame-of-ref | 25 |
| 2.6 | f-o-r-deiktisch | 26 |
| 2.7 | f-o-r-intrinsic | 26 |
| 2.8 | f-o-r-extrinsic | 27 |
| 2.9 | hermann-ebene | 29 |
| 2.10 | hermann-deiktisch | 31 |
| 2.11 | linearisierung | 46 |
| 2.12 | navigator | 51 |
| 2.13 | spat-struct-gopal | 51 |
| 2.14 | habel-system | 58 |
| 2.15 | coucelis-system | 60 |
| 2.16 | rau-system | 62 |
| 2.17 | kleist-system | 63 |
| 2.18 | carstensen-system | 65 |
| 3.1 | karte-campus | 76 |
| 3.2 | fragment-haeufig | 77 |
| 3.3 | fragment-komb-haeuf | 78 |
| 3.4 | zeitrahmen | 78 |
| 3.5 | kat-pro-zeitrahmen | 80 |
| 3.6 | kat-komb-pro-zrahmen | 80 |
| 3.7 | exp-4-crossing | 84 |
| 3.8 | color-intens | 86 |
| 3.9 | right-wrong-side | 87 |
| 3.10 | height-intens | 88 |
| 3.11 | vgl-hoche-farbe | 89 |
| 4.1 | modules | 93 |
| 4.2 | env2map-mapping | 94 |
| 4.3 | pfad-ep | 96 |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>spatial-module</td>
<td>99</td>
</tr>
<tr>
<td>4.5</td>
<td>anchoring</td>
<td>101</td>
</tr>
<tr>
<td>4.6</td>
<td>3d-campus-sgi</td>
<td>104</td>
</tr>
<tr>
<td>4.7</td>
<td>volllst-formales-modell</td>
<td>105</td>
</tr>
<tr>
<td>4.8</td>
<td>time-intervals</td>
<td>106</td>
</tr>
<tr>
<td>4.9</td>
<td>time-optimization</td>
<td>108</td>
</tr>
<tr>
<td>4.10</td>
<td>kontrollfluß-module</td>
<td>109</td>
</tr>
<tr>
<td>4.11</td>
<td>pfadintegriationsphase</td>
<td>111</td>
</tr>
<tr>
<td>4.12</td>
<td>objekti integrationsphase</td>
<td>112</td>
</tr>
<tr>
<td>4.13</td>
<td>WBgenerierungsphase</td>
<td>113</td>
</tr>
<tr>
<td>4.14</td>
<td>phases-flow</td>
<td>114</td>
</tr>
<tr>
<td>4.15</td>
<td>kontroll-modul</td>
<td>115</td>
</tr>
<tr>
<td>4.16</td>
<td>inkr-pfadsuche</td>
<td>118</td>
</tr>
<tr>
<td>4.17</td>
<td>corridor</td>
<td>119</td>
</tr>
<tr>
<td>4.18</td>
<td>entscheidungs-pkt</td>
<td>122</td>
</tr>
<tr>
<td>4.19</td>
<td>freier-platz</td>
<td>123</td>
</tr>
<tr>
<td>4.20</td>
<td>kreisverkehr</td>
<td>124</td>
</tr>
<tr>
<td>4.21</td>
<td>lineare-i-rek-merkmale</td>
<td>126</td>
</tr>
<tr>
<td>4.22</td>
<td>ausw-saliener-obj</td>
<td>128</td>
</tr>
<tr>
<td>4.23</td>
<td>projection</td>
<td>128</td>
</tr>
<tr>
<td>4.24</td>
<td>spatial-focus</td>
<td>129</td>
</tr>
<tr>
<td>4.25</td>
<td>focus-salience</td>
<td>130</td>
</tr>
<tr>
<td>4.26</td>
<td>sit2Raster-locs</td>
<td>131</td>
</tr>
<tr>
<td>4.27</td>
<td>raster-error</td>
<td>131</td>
</tr>
<tr>
<td>4.28</td>
<td>cie-diagram</td>
<td>133</td>
</tr>
<tr>
<td>4.29</td>
<td>color-diffs</td>
<td>133</td>
</tr>
<tr>
<td>4.30</td>
<td>height-map-serra-sgi</td>
<td>138</td>
</tr>
<tr>
<td>4.31</td>
<td>width-map-serra-sgi</td>
<td>139</td>
</tr>
<tr>
<td>4.32</td>
<td>global-map-serra-sgi</td>
<td>139</td>
</tr>
<tr>
<td>4.33</td>
<td>sal-obj-1</td>
<td>140</td>
</tr>
<tr>
<td>4.34</td>
<td>dynamic-sit</td>
<td>142</td>
</tr>
<tr>
<td>4.35</td>
<td>prs-orientation</td>
<td>144</td>
</tr>
<tr>
<td>4.36</td>
<td>prim-ref-sys</td>
<td>144</td>
</tr>
<tr>
<td>4.37</td>
<td>srs-orientation</td>
<td>145</td>
</tr>
<tr>
<td>4.38</td>
<td>sek-ref-sys</td>
<td>145</td>
</tr>
<tr>
<td>4.39</td>
<td>sek-ref-sys2</td>
<td>146</td>
</tr>
<tr>
<td>4.40</td>
<td>vrs-orientation</td>
<td>147</td>
</tr>
<tr>
<td>4.41</td>
<td>virt-ref-sys</td>
<td>148</td>
</tr>
<tr>
<td>4.42</td>
<td>( t_{or} - tilted - A )</td>
<td>148</td>
</tr>
<tr>
<td>4.43</td>
<td>vrs-auto</td>
<td>149</td>
</tr>
<tr>
<td>4.44</td>
<td>@</td>
<td>152</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.45</td>
<td>right-of-deikt</td>
<td>153</td>
</tr>
<tr>
<td>4.46</td>
<td>eval-intrinsic</td>
<td>154</td>
</tr>
<tr>
<td>4.47</td>
<td>gen-rkb</td>
<td>156</td>
</tr>
<tr>
<td>4.48</td>
<td>integr-path-elem</td>
<td>157</td>
</tr>
<tr>
<td>4.49</td>
<td>integrate-path-segment</td>
<td>157</td>
</tr>
<tr>
<td>4.50</td>
<td>topo-intrin-all</td>
<td>158</td>
</tr>
<tr>
<td>4.51</td>
<td>path-intention-pattern</td>
<td>160</td>
</tr>
<tr>
<td>4.52</td>
<td>proto-conf-intent</td>
<td>160</td>
</tr>
<tr>
<td>4.53</td>
<td>select-srs</td>
<td>162</td>
</tr>
<tr>
<td>4.54</td>
<td>integr-lm</td>
<td>162</td>
</tr>
<tr>
<td>4.55</td>
<td>topo-intrin-all-lm</td>
<td>164</td>
</tr>
<tr>
<td>4.56</td>
<td>loc-of-lms</td>
<td>165</td>
</tr>
<tr>
<td>4.57</td>
<td>appl-of-lm</td>
<td>166</td>
</tr>
<tr>
<td>4.58</td>
<td>anchoring-prs-in-nav</td>
<td>167</td>
</tr>
<tr>
<td>4.59</td>
<td>animation-seq-move</td>
<td>167</td>
</tr>
<tr>
<td>4.60</td>
<td>animation-seq-rot</td>
<td>168</td>
</tr>
<tr>
<td>4.61</td>
<td>message-structure</td>
<td>170</td>
</tr>
<tr>
<td>4.62</td>
<td>mesg-structure-partial</td>
<td>173</td>
</tr>
<tr>
<td>4.63</td>
<td>lm-not-refo</td>
<td>176</td>
</tr>
<tr>
<td>4.64</td>
<td>desc-lm</td>
<td>177</td>
</tr>
<tr>
<td>4.65</td>
<td>pagm</td>
<td>178</td>
</tr>
<tr>
<td>4.66</td>
<td>inh-kat-komb</td>
<td>183</td>
</tr>
<tr>
<td>4.67</td>
<td>gui-moses</td>
<td>186</td>
</tr>
<tr>
<td>4.68</td>
<td>gui-campus</td>
<td>188</td>
</tr>
<tr>
<td>4.69</td>
<td>fuss-increm-2jams</td>
<td>189</td>
</tr>
<tr>
<td>4.70</td>
<td>lfl</td>
<td>189</td>
</tr>
<tr>
<td>4.71</td>
<td>lfl0</td>
<td>190</td>
</tr>
<tr>
<td>4.72</td>
<td>gui-moses-view</td>
<td>190</td>
</tr>
<tr>
<td>4.73</td>
<td>gui-moses-speak</td>
<td>195</td>
</tr>
<tr>
<td>4.74</td>
<td>compl-descr-foot.pic</td>
<td>196</td>
</tr>
<tr>
<td>4.75</td>
<td>compl-descr-car.pic</td>
<td>197</td>
</tr>
</tbody>
</table>
Kapitel 1

Einführung


1.1 Problemstellung

Um sich adaptiv und nutzbringend in seiner Alltagsumgebung verhalten zu können, ist es für den Menschen entscheidend, aus der Umgebung die Information herauszufiltern, die bezüglich seiner Intentionen wichtig erscheint. Von einer internen Perspektive aus gesehen, bedingen Kapazitätsbeschränkungen des menschlichen kognitiven Systems, dass er sich bei der Auswahl stark einschränken muss. Neben diesen internen Beschränkungen sind externe Beschränkungen zu beachten, die durch die Umgebung oder das Verhalten des Menschen in der Umgebung vorgegeben sind. Beispielsweise gibt die Bewegung eines Sprechers relativ zur nächsten Kreuzung vor, inwieweit sein Verhalten zeitlich und räumlich eingeschränkt ist. Aus dem jeweiligen Verhalten leitet der Sprecher die Beschränkungen ab, verinnerlicht diese und macht sie zur Grundlage seiner Verhaltensfindung. Diese Beschränkungen bedingen, dass der Mensch in möglichst effizienter Weise Information auswählt, um in effektiver Weise agieren zu können.


In Tabelle 1.1 sind zum Vergleich eine vollständige und eine inkrementelle Wegbeschreibung gegenübergestellt. Sie beschreiben jeweils denselben Pfad vom Gebäude der Informatik zur Hauptbibliothek der Universität des Saarlandes. Auffällig ist, daß die vollständige im Vergleich zur inkrementellen Wegbeschreibung wesentlich kürzer und ungenauer ist. Ferner werden vielmehr große Gebäude referenziert.

1.2 Einordnung der Arbeit

1.2. EINORDNUNG DER ARBEIT

Also, dann gehe unter der Unterführung durch, an der Mathematik vorbei,
auf der rechten Seite kommt dann ein freier Platz,
gehe über den Platz geradeaus die Treppe hoch,
dann kommst du an eine Straßensache, wo du so links gehest,
dann siehst du den EDEKA-Markt,
dann gehts rechts an diesem EDEKA vorbei,
dann ist da auch die Bayerische Vereinsbank,
dann gehest du hoch,
dann links abbiegen,
dann bei der Post rechts abbiegen und dann links die Treppen hoch,
dann bist du auf einem großen freien Platz,

dort nach rechts,
und dann sieht man auch schon den hohen Büchereiturm,
Aber bei Bock und Seip solltest du am besten nochmal nachfragen

Hier jetzt geradeaus, die Treppe runter, über die Straße dann da die Treppe wieder hoch.


Hier jetzt wieder ein Stück geradeaus ... an dem Gebäude. Jetzt kommen wieder Treppen ... hoch.

Hier über den Weg.

Und dann die großen Treppen da vorne wieder hoch.

Geradeaus hier. Und da vorne bei der Wegkreuzung vor dem Glaskontainer oder Altpapierkonto-
tainer ... dann links ... OK, jetzt links hier.

Hier geradeaus ein Stück und oben an der Kreuzung dann wieder rechts. So jetzt hier rechts.

Geradeaus. Treppe hoch. Und hier vorne jetzt an der Kreuzung links.

Hier jetzt geradeaus. Und vor dem Musikbau da vorne auf der linken Seite gehts rechts die Treppe
hoch. Also zwischen Post und Musikgebäude. So da jetzt rechts rüber, da wo die Leute herkommen.

Hier jetzt links die Treppe hoch.

Jetzt, hier rechts.

Jetzt einfach geradeaus. Und da vorne ist auch schon die Bibliothek. Hier jetzt einfach über die 
Straße über ... über den Zebrastreifen und da wo Fahrräder und die Motorräder stehen, da ist der 
Eingang. Da gehts rein.

Tabelle 1.1: Beispiele für eine vollständige und eine inkrementelle Wegbeschreibung

nerieren. Navigationshilfesysteme werden bereits punktuell im Automobilverkehr eingesetzt. Sie
besitzen jedoch einen viel größeren Einsatzbereich, wie nicht zuletzt zu Rehabilitationszwecken
bei neurophysiologisch oder perzeptive geschädigten Personen. Beide Perspektiven lassen sich
unter der Prämisse zusammenfassen, daß ein Navigationshilfesystem um so nutzbringender ist, 
je mehr es sich wie ein Mensch verhält. Dies aber kann nur dann verwirklicht werden, wenn
man die Prozesse und Repräsentationen, die das Verhalten des Menschen determinieren, besser
versteht.

Die vorliegende Arbeit ist im Kontext der Künstlichen Intelligenz (KI) Forschung erstellt worden. 
Eines der Arbeitsgebiete befaßt sich mit der Interaktion von visueller Wahrnehmung, räumlichem
Wissen und Sprache. Besonders in Deutschland ist die Forschung in dieser Richtung sowohl aus
kognitionswissenschaftlicher als auch ingenieurwissenschaftlicher Sicht vorangetrieben worden. 
Bereits mit dem System HAM-RPM [vgl. [WJH78]] und nachfolgend mit SWYSS (vgl. [HS84])
wurde ein einfaches Modell der natürlichsprachlichen Beschreibung statischer räumlicher Infor-
mation vorgestellt, wobei insbesondere auch Aspekte der Vagheit sprachlicher Raumbeschreibungen ihre Berücksichtigung fanden. Während in SWYSS eine Seitenansicht eines Zimmers als Ausgangsszene behandelt wurde, stand in den Systemen GEOSYS (vgl. [FKR84]), LANDSCAN (vgl. [BJKZ88]) und LEI (vgl. [CMC94]) der Zugriff auf geographische Information und die damit verbundenen räumlichen Beziehungen im Vordergrund. Die Aufgabenstellung in den Systemen WINTOP (vgl. [OMT94]) und VIENA (vgl. [WC94]) ist es, geometrische Szenenmodelle in einem Graphiksystem durch sprachliche Interaktion inkrementell zu verfeinern. Neben rein softwarebasierten Ansätzen wird in einigen Arbeiten untersucht, wie die Integration von visuellen Daten und Sprache zur Steuerung von Robotern verwendet werden kann (vgl. [LFMM94], [HMRS95], [SGH+ 94]). Das Projekt LILOG-R ([HR91], [Pri93]) untersuchte vor allem Repräsentationen, die das Verstehen von Lokalisierungsaussprüchen bezüglich Situationen, zu denen der Rezipient keinen unmittelbaren perzeptuellen Zugang besaß, modellieren. Ein Überblick zu den aktuellen Forschungsaktivitäten im Bereich der Integration von perzeptuellen und sprachlichen Fähigkeiten findet sich in verschiedenen Aufsatzsammlungen (vgl. [McK94b], [McK94a]).

Im Kontext der Integration von bildverstehenden und sprachverstehenden Systemen ist die Domäne der Wegbeschreibungen, aufgrund der Abgrenztheit des Phänomenbereichs, aus linguistischer, psychologischer und KI-Sicht bezüglich des dynamischen Zusammenspiels von visueller Perzeption, Raumkognition und Sprache häufig untersucht worden. In linguistischen Arbeiten sind vor allem syntaktische und semantische Aspekte von sprachlichen Wegbeschreibungen analysiert worden (z.B. [Kle82], [WR82], [Hab87], [MMP+88], [HCR90]). Psychologisch motivierte Arbeiten haben anhand von Wegbeschreibungen besonders räumliches Wissen untersucht (z.B. [HM71], [DS73], [Tw92]). Beide Aspekte spielen in algorithmischen Modellen eine wesentliche Rolle (z.B. [Kui78], [Dav86], [GKS89], [LZ89]).

Die hier vorgestellte Arbeit ist im Kontext der Projekte VITRA1 und REAL2 entstanden. In VITRA ist es gelungen, erste Erkenntnisse über die Durchschaltung visueller Reaktionen zu semantischen Beschreibungen zu gewinnen ([HSA+89]). Bezüglich der Domäne des Fußballspiels ist im Rahmen von VITRA ein System namens SOCCER entwickelt und implementiert worden, welches visuelle Daten von realen Fußballszenen verbal beschreibt (vgl. [HRA89], [HW94], [Sch94]). Weiterhin wird in einer Konstruktionsdatenbank untersucht, wie ein Roboter mit zwei Effektoren natürlichsprachlich gesteuert werden kann (vgl. [SL96]).

In Zusammenarbeit mit Forschern der Visuellen Datenverarbeitungsgruppe am Fraunhofer Institut (ITB) in Karlsruhe sind im Rahmen des Sonderforschungsbereichs 314 „Wissensbasierte Systeme“ Techniken untersucht worden, mit denen Videodaten von Fußballspiel- und Verkehrsszenen analysiert und verbal beschrieben werden können. Als Schnittstelle zwischen visuellen Daten und sprachlichen Beschreibungen werden geometrische Szenenbeschreibungen verwendet, die zeitlich ausgedehnte Szenen über die Festlegung von Lokationen relativ zu einem Koordinatensystem integrieren (vgl. [Neu86]). Auf der Basis der geometrischen Szenenbeschreibung wird die statische und dynamische, räumliche Struktur der Szene analysiert. Einerseits werden dazu statische, räumliche Lagebeziehungen zwischen Objekten ([And88]) und andererseits Ereignisse erkannt (vgl. [HR88], [Her96]). Neben diesen rein räumlichen Analysen, können situationsbezogene Intentionen abgeleitet werden (vgl. [RS92]). Räumliche Beschreibungen erzeugen beim Rezipienten räumliche Vorstellungen. Perzipiert der Rezipient den beschriebenen Sachverhalt, so kann er/sie diese räumliche Vorstellung sofort verifizieren. In Situationen, wie beispielsweise bei einer Radioreportage eines Fußballspiels, ist dies nicht möglich. In einem solchen Fall ist eine solche Beschreibung zu generieren, die es dem Rezipienten erlaubt, eine geeignete

1VITRA steht für „visual translator“, welches das Teilprojekt N2 des SFB 314 (Wissensbasierte Systeme) war (vgl. [?], [HBC+86], [Wal94]).
2REAL steht für „Ressourcenadaptive Objektlokalisierung“, welches das Teilprojekt A4 des SFB378 (Ressourcenadaptive cognitive Prozesse) ist (vgl. [TLS+95]).
1.3. ZIELSETZUNG UND METHODOLOGIE

mentale räumliche Repräsentation der Szene aufzubauen. Bevor der Sprecher eine Beschreibung präsentiert, antizipiert er durch Konstruktion eines „mentalen Bildes“ die Vorstellung beim Rezipienten, die durch die Beschreibung erzeugt wird, und paßt seine Beschreibung gegebenenfalls an (vgl. [WJH78], [Sch94]). Ein Berechnungsmodell einer solchen Antizipationsrückkopplungsschleife ist in SOCCER integriert (vgl. [Sch94],[SS93], [BS95]). Im Kontext von VITRA ist mit dem System CITYGUIDE ein erstes Modell für die Generierung von vollständigen Wegbeschreibungen entwickelt worden (vgl. [Müll88]). Als Weiterentwicklung der in VITRA entwickelten Konzepte entstand die Fragestellung nach einem Modell zur Generierung inkrementeller Wegbeschreibungen. Die Zielsetzung und die verwendete Methodologie sollen nun kurz vorgestellt werden.

1.3 Zielsetzung und Methodologie


1. Zeitabhängige Generierung inkrementeller Wegbeschreibungen
   Wie die empirischen Untersuchungen ferner zeigen, wird eine inkrementelle Wegbeschreibung an zeitliche und räumliche Ressourcenbeschränkungen adaptiert, die durch die Bewegung des Sprechers relativ zur Umgebung gegeben sind. Die Komplexität einer Beschreibung hängt von dem Zeitraum ab, den der Sprecher bis zum Erreichen des nächsten Verzweigungspunktes zur Verfügung hat. Entsprechend dieser Beschränkungen werden geeignete Inhaltsstrukturen ausgewählt, Oberflächenstrukturen transformiert und zu geeigneten Zeitpunkten präsentiert.

2. Auswahl visuell salienter Objekte in simulierten 3D-Umgebungen
   Die visuelle Salienz eines Objektes hängt von seinen visuellen Merkmalen und dem Kontext, der durch die gesamte Situation gegeben ist, ab. Visuelle Saliencen lassen sich bezüglich einzelner Merkmalsklassen bestimmen, die, relativ zur Umgebung gewichtet, zu einem Salienzwert eines Objektes integriert werden.

3. Inkrementelle heuristische Suche von Pfaden in topographischen Netzen
   Ausgehend davon, daß der Sprecher kein räumliches Wissen bezüglich der Umgebung besitzt, wird eine topographieerhaltende Kartenrepräsentation verwendet, um geeignete Pfadelemente (Pfadabschnitte und Kreuzungen) inkrementell unter Einbeziehung von Wissen zu bestimmen.

4. Etablierung geeigneter räumlicher Referenzsysteme
   In einer Umgebung benötigt ein Sprecher eine Orientierung, auf dessen Basis er die räumliche Lage von Objekten zu ihm selbst oder relativ zwischen Objekten bestimmen kann. Zu diesem Zweck sind verschiedene räumliche Referenzsysteme zu etablieren, die durch
die Bewegung oder Blickrichtung des Sprechers bzw. durch intrinsische Eigenschaften der Objekte gegeben sind.

5. Etablierung einer *minimalen* Anzahl räumlicher Relationen
Unter der Annahme, daß der Mensch aus Ökonomie- und Effizienzgründen versucht, eine möglichst geringe Anzahl an räumlichem Wissen aufzubauen, wird analysiert, zwischen welchen Objekten räumliche Relationen etabliert werden müssen, um Wegbeschreibungen generieren zu können. Hierzu werden gradierte räumliche Relationen verwendet.

6. Schemabasierte Ableitung *pfadgeleiteter Intentionen*

7. Situationsadaptive Auswahl sprachlicher *Inhaltsstrukturen*
Wie die in Kapitel 3 vorgestellten empirischen Untersuchungen zeigen, werden deutschsprachige inkrementelle Wegbeschreibungen meist durch schematische Kombination einzelner inhaltlicher Teile aufgebaut. Diese Teile (Inhaltskategorien) beziehen sich auf die Beschreibung einer Lokation, einer Richtung, eines Zeitpunktes oder einer Aktion und werden zu *Inhaltsstrukturen* zusammengesetzt.

8. Schemabasierte Auswahl räumlicher *Präpositionen* und *Adverbia*
Räumliche Information wird in Wegbeschreibungen vor allem durch Präpositionen und Adverbia ausgedrückt. Es wird auf der Basis des dynamisch etablierten räumlichen Wissens ein Verfahren zu Lexikalisation vorgestellt.

Da bisher kein Korpus inkrementeller Wegbeschreibungen frei verfügbar ist, mußte als Teil dieser Arbeit ein entsprechender Korpus durch eine Felduntersuchung aufgebaut werden.\(^3\) Aus diesem Korpus sind inhaltliche Strukturen und zeitliche Abhängigkeiten abgeleitet worden. In einer zweiten empirischen Untersuchung ist bezüglich einer computeranimierten 3D-Umgebung analysiert worden, in welcher Weise die Variation visueller Merkmale und pfadbasierter Intentionen die Auswahl von Landmarken, die meist als Referenzobjekte verwendet werden, beeinflußt.


\(^3\)Am LIMSI in Orsay ist in einem Fahrer-Beifahrer-Szenario ein 64-stündiger Korpus französischer inkrementeller Wegbeschreibungen aufgebaut worden (vgl. [BD93]). Da es sich hierbei um eine Auftragsarbeit für einen Automobilkonzern handelt, ist der Korpus nicht frei verfügbar.


1.4 Gliederung

Die weiteren Kapitel sind entsprechend der verwendeten Methodologie Empirie-Modellbildung-Implementation aufgebaut.

In Kapitel 2 wird zur Theorie behandelt, um einen Überblick über das gesamte Problemfeld der Raumkognition im allgemeinen und das von Wegbeschreibungen im speziellen aufzuzeigen und um die Terminologie einzuführen. Hierbei wird der aktuelle Forschungsstand hauptsächlich aus den Bereichen der Kognitiven Psychologie, der ökologischen Psychologie, der Anthropo-Geographie, der Psycho- und Computerlinguistik und der künstlichen Intelligenz systematisch im Hinblick auf die oben dargestellten Kernpunkte ausgewertet.

Daran anschließend werden in Kapitel 3 empirische Untersuchungen und deren Ergebnisse vorgestellt, die im Rahmen dieser Arbeit durchgeführt worden sind. Dieses Kapitel ist in zwei empirische Untersuchungen unterteilt: ein Feldexperiment und ein Laborexperiment.

In Kapitel 4 wird das formale Modell beschrieben, welches sich in den Rahmen der Theorie und
der Empirie einbettet. Es unterteilt sich in eine Vorstellung der Module und Funktionen und abschließend der Implementation. Das formale Modell, welches durch die Module und Funktionen beschrieben ist, wird als generischer Rahmen verstanden, welcher in der Implementation MOSES eine mögliche Realisierung gefunden hat.

Abschließend wird der Wert, den dieses Modell aus einer Ingenieursperspektive besitzt, in Kapitel 5 zusammen mit den offenen Fragen und einem Blick in die Zukunft diskutiert.
Kapitel 2

Theorien zur Raumkognition


In welcher Weise räumliches Wissen abgeleitet, repräsentiert, evaluiert und genutzt wird, ist der Ansatzpunkt für eine Vielzahl von Forschungsrichtungen. Durch neurophysiologische Stu-


“...A theory of cognition should have two aspects. On the one hand, there should be a formal characterization of the functions or transformations performed by the cognitive system; on the other hand, there should be a description of the psychological mechanisms whereby such a system could be realized. Although there is considerable information about neural mechanisms underlying afferent processes, and a picture of some of the mechanisms underlying sensation and perception is beginning to emerge, the possibility of saying anything definitive about the conceptual mechanisms of the nervous system is sufficiently remote at the present time to justify omitting this aspect of a comprehensive theory from our considerations. At the level of sensation, however, enough work has been done, both in the formulation of laws relating sensory attributes to physical stimulation and in the analysis of processes in the sensory receptors and nerves, that we might expect to find the beginnings of a comprehensive theory.” (vgl. [MJL76, S. 12])


Der Mensch verwendet räumliches Wissen, um sich in Umgebungen zu orientieren und intentionsbasiert zu agieren. Eine solche phänomenologische Beschreibung gibt jedoch keinerlei Auskunft über die Verarbeitungsmechanismen und Repräsentationen.


Bei der Untersuchung raumkognitiver Fähigkeiten wird vielfach eine Unterscheidung nach dem Verhältnis eines externen Weltraumauschnitts zum Akteur vorgenommen. Ist der betrachtete Weltraumauschnitt größer als das, was der Agent von einer Lokation aus wahrnehmen kann, so spricht man von large scale space (vgl. [Kui78]). Im weiteren Verlauf wird dies als Umgebung bezeichnet.

Der Weltraumauschnitt, den der Agent von einer Lokation aus vollständig perzipieren kann, heißt üblicherweise small scale space und wird im folgenden als Situation bezeichnet. Situationen sind demzufolge Ausschnitte aus Umgebungen.


Eine Aufgabe raumkognitiver Leistungen ist es, aus Wahrnehmungsreizen extrahiertes, räumliches Wissen zu etablieren, miteinander zu verknüpfen, zu speichern und wieder zugänglich zu machen. Solcherart bestimmtes räumliches Wissen bezieht sich auf räumliche Lokationen, Ver teilungen, Muster, Zusammenhänge, Konfigurationen und andere Entitäten, die einem bei der Bestimmung eines Pfades, der Aktivierung von Bewegungen, aber auch zur Beschreibung von räumlichen Sachverhalten dienen.

1Soweit es nicht anders vorgegeben wird, bezieht sich alles Folgende auf reale, d.h. nach physikalisch-mathematischen Gesetzen formalisierbare Umgebungen. Da u.a. aus ontogenetischen Gründen davon ausgegangen wird, daß der Umgang mit realem Raum die Basis für den Aufbau und die Verfeinerung von raumbezogenen Kon zepten und Prozessen insgesamt ist, kann diese Annahme ohne Beschränkung der Allgemeinheit gemacht werden.
Um aus einer Situation Information auswählen zu können, verfügt der Mensch über Fokussierungsmechanismen. Ein solcher Mechanismus ist der der Aufmerksamkeit. In Bezug auf Aufmerksamkeit bezüglich visueller Wahrnehmung gibt es verschiedene Ansätze. Aus dem anfänglichen Scheinwerfer-Modell (vgl. [EE74]) entstanden die Zoomlinsen- (vgl. [ES86]) und die Gradienten-Theorie (vgl. [LB89]). Zu diesen Theorien sind zahlreiche Experimente zur räumbezogenen und zur objektbezogenen Aufmerksamkeit durchgeführt worden (vgl. [Dun84], [KH81], [KTC92], [KJ91]). Die meisten Theorien zur visuellen Aufmerksamkeit nehmen diskrete Lokationen als Basis für Aufmerksamkeitsfunktionen an (z.B. [ES86], [LB89], [PC84], [TG80], [van92]). In diesem Kontext entwickelte Anne Treisman die lokationsbasierte feature integration theory, die auf einer merkmalsbasierten Indizierung von Lokationen beruht (vgl. [TG80], [Tre88], [Tre93] und Abbildung 2.1).

Abbildung 2.1: Feature Integration Theory (nach [Tre88])

In dieser Theorie nimmt Treisman an, daß verschiedene sensorische Merkmale, wie „Farbe“, „Orientierung“, „Größe“ und „Bewegungsrichtung“ in spezialisierten Modulen parallel und automatisch verarbeitet werden, während Objekte erst danach aus diesen Merkmalen mittels fokussierter Aufmerksamkeit identifiziert werden. Merkmale³ werden automatisch und räumlich parallel verarbeitet, ohne daß dazu fokussierte Aufmerksamkeit notwendig ist (vgl. [Tre88, S. 203]). Eine solche Hypothese basiert auf anatomischen und physiologischen Befunden, in denen auf visuelle Kortexareale hingewiesen wird, die auf verschiedene Aufgaben spezialisiert sind (vgl. [Cow79], [MN87], [vM88], [Zek81]). Grundlage für die feature integration theory ist eine zweidimensionale, gerasterte master map of locations (siehe Abbildung 2.1). Von dieser zweidimensionalen Repräsentation werden für jeden Merkmalsaspekt durch Projektionen Merkmalskarten abgeleitet. Wenn Objekte gesucht werden, wird ein Aufmerksamkeitsbereich auf der master map etabliert und dieser Bereich an die Merkmalskarten weitergeleitet. In den einzelnen Merkmalskarten sind alleinig die Merkmalswerte, aber nicht deren Lokationen repräsentiert. Merkmale

²Für einen Überblick siehe [Log96].
³Treisman läßt die Frage nach der Definition von Merkmalen für empirisch, die durch Konvergenzuntersuchungen beantwortbar ist (vgl. [Tre88, S. 230]).


Bezogen auf raumkognitive Prozesse wird angenommen, daß Aufmerksamkeit notwendig ist, um räumliche Relationen zu berechnen (z.B. [EZ83], [Log95]). Danach werden keine räumlichen Relationen zwischen Objekten berechnet, die nicht im Aufmerksamkeitsbereich liegen. Aufmerksamkeit ist dazu notwendig, aus der prinzipiell unendlichen Menge von möglichen Relationen geeignete auszuwählen. Zwischen Entitäten im Aufmerksamkeitsbereich, werden räumliche Relationen solange nicht evaluiert, wie dies nicht intendiert ist (vgl. [Log95, S. 163]). Dies stellt eine derartige starke Anforderungen an die Funktionen zur Berechnung räumlicher Relationen. andererseits bewirken diese Beschränkungen eine Reduzierung der Komplexität.

Im alltäglichen Leben werden Beschreibungen räumlicher Sachverhalte vielfach verwendet. Dabei wird zwischen realen, perzepierbaren und imaginierten Räumen unterschieden. Ausgehend von der ontogenetischen Entwicklung eines Menschen ist der reale, geographisch-physikalische Raum des Alltags als prototypisch anzusehen. Von diesem leiten Kinder raumbezogene Konzepte ab, die sie später auch auf abstrakte Domänen anwenden (vgl. [Pie67]). Obschon in diesem Kontext

---

Metaphorisch wird das Konzept des Raumes auch auf nicht-räumliche Domänen bezogen (vgl. [LJ80]).
KAPITEL 2. THEORIEN ZUR RAUMKOGNITION

eine Vielzahl von Einzelphänomenen untersucht worden sind, ist der konzeptionelle Rahmen der Raumkognition nach wie vor erst schemenhaft zu erkennen.


![Abbildung 2.2: Ein Beispiel für die Verankerung sprachlicher Raumausdrücke in visuellen Daten](image)

Dadurch sind sprachliche Äußerungen über räumliche Sachverhalte im räumlichen Wissen des Sprechers verankert. Dieses Wissen wiederum ist in der Information verankert, die dem Spre-

Ein anderer Vorschlag zur Konstruktion einer Semantiktheorie ist die *Situationssemantik* von Barwise und Perry (vgl. [BP83]). Sie geben einen allgemeinen Formalismus an, in dem sich relationale Beziehungen zwischen Entitäten in einer Situation und ihre Wirkung auf sprachliche Aussagen beschreiben lassen. Die Semantik einer Aussage ist dadurch nicht über Wahrheitsbedingungen oder Mengen möglicher Welten bestimmt, sondern über *Situationen* und *Beschränkungen* zwischen Situationen. Beschränkungen können formale aber auch naturgesetzliche Regelmäßigkeiten repräsentieren, wodurch Beschränkungen Bedeutung konstituieren (vgl. [Pin92]). Nomina werden durch Referenzrelationen in Objekten der Situation verankert.6 Dadurch definiert sich beispielsweise die Bedeutung eines Bezeichners $\beta$ im Satz $\phi$ durch die Relation zwischen der Aussage, in der $\beta$ vorkommt, und den Individuen, auf die $\beta$ bezogen ist. Bisher konnte jedoch keine auch nur annähernd vollständige Theorie auf der Basis der Situationssemantik vorgestellt werden (vgl. [Pin92]).

Basierend auf einer geeigneten Semantiktheorie wird bei der Generierung einer Aussage konzeptuell unterschieden *was* und *wie* etwas gesagt werden soll (vgl. [Tho77]). Als Konzeptualisierer und *Formulierer* bezeichnet, ist diese Dichotomie aus psycholinguistischer Sicht untersucht und belegt worden (vgl. [KH87], [Lev89]). Eine ähnliche Unterscheidung wird in multimodalen Generierungssystemen verwendet (z.B. [FM89], [WAF+92], [Wah94]). Der Konzeptualisierer unterteilt sich nach Kempen und Hoenkamp in die Phase der *Makroplanung*, in welcher Sprechakte bestimmt werden, und in die Phase der *Mikroplanung*, in der präverbalen Bedeutungsstrukturen festgelegt werden. Die *grammatikalische* und *phonologische* Enkodierung erfolgt im *Formulier* (vgl. [Lev89]). Dazu wird die präverbalen Bedeutungsstruktur in *Oberflächenstrukturen* transformiert, welche dann in *Artikulationspläne* umgesetzt werden (vgl. [KH87]). In einer nachfolgenden Komponente, dem sogenannten *Artikulator*, werden Artikulationspläne ausgeführt (vgl. [Lev89]). Da ein Sprecher seine Aussagen selbst mitkontrollieren kann, nimmt Levelt eine *Sprachverstehenskomponente* an. In dieser analysiert der Sprecher seine *interne* Sprache und modifiziert gegebenenfalls die gesamte Äußerung (vgl. [De180], [Lev83]). Eine präverbalen Bedeutungsstruktur wird als eine propositional dargestellte semantische Repräsentation angesehen (vgl. [Lev89]).

In den meisten Sprachgenerierungssystemen sind präverbalen Bedeutungsstrukturen an die jeweilige Domäne angepaßt. Levelt schlägt eine hierarchische *Funktor/Argument*-Struktur vor, die aus dem Zusammenspiel verschiedener kognitiver Systeme entsteht (vgl. [Lev89, S. 78]). Dieser Ansatz basiert auf Jackendoffs *konzeptueller Semantik*, die eine Zwischenstufe zwischen allgemeiner Semantiktheorie und Sprachtheorie darstellt (vgl. [Jac83], [Jac87], [Jac90]). Levelt unterscheidet zehn Kategorien (vgl. Tabelle 2.1), aus denen sich präverbalen Bedeutungsstrukturen zusammensetzen (ähnlich zu semantischen Netzen (vgl. [B885]) und propositionalen Netzen (vgl. [And83])). Nur *EVENT* und *STATE* repräsentieren dabei vollständige Propositionen (vgl. [Lev89, S. 78]). Alle anderen sind elliptisch oder aber mit anderen Kategorien zu vollständigen zusammensetzen. Beispielsweise können *THING* und *PLACE* die Dekomposition einer Kate-

---

6Die formale Repräsentation ist wie folgt strukturiert: d.c[φ]e mit der Diskurssituation d, der Referenzrelation c, Satz $\phi$ und der Situation e, auf die sich die Äußerung bezieht.
Kategorie | Beschreibung
--- | ---
PERSON | Referenz auf eine Person
THING | Referenz auf ein Objekt
EVENT | Referenz auf ein Ereignis
ACTION | Referenz auf eine Aktion
STATE | Referenz auf eine Zustandsbeschreibung
TIME | Temporale Referenz
PLACE | Referenz auf eine Lokation
DIRECTION | Referenz auf eine Richtung
ATTRIBUTE | Referenz auf ein Attribut eines Objektes
MANNER | Referenz auf einen Modus einer Aktion

Tabelle 2.1: Kategorien der Bedeutungsstrukturen nach Levelt (vgl. [Lev89]).

gorie STATE sein.


Kognitive Beschränkungen des Sprechers, aber auch die vom Sprecher beim Rezipienten vermuteten, erfordern, daß der Inhalt einer Nachricht fokussiert wird (vgl. [GS86], [Her83]). Inhaltsliche Sprünge sind dabei zu vermeiden oder so einzusetzen, daß sie vom Hörer nachvollziehbar sind. Wird ein wahrnehmbarer Sachverhalt beschrieben, so wird die Fokussierung teilweise durch diesen übernommen. Möchte der Sprecher bestimmte Aspekte eines Sachverhaltes hervorheben, so muß die Aufmerksamkeit des Hörers durch Beschreibungen⁶ gelenkt werden.

Ist der Inhalt der intendierten Nachricht festgelegt, so wählt der Sprecher geeignete Informationsseinheiten aus, von denen er erwartet, daß sie beim Rezipienten zum Verständnis der Aussage beitragen werden. Präferiert der Rezipient minimale Beschreibungen und leistet der Sprecher diesem Wunsch Folge, so wird der Sprecher eine weniger detaillierte Beschreibung generieren, als wenn der Sprecher annimmt, daß der Rezipient eine ausführliche Beschreibung bevorzugt. Als allgemeine Richtlinie können die Griechen Konversationsmaximen verwendet werden, die vom Sprecher verlangen, daß er nur soviel beschreibt wie nötig und so wenig wie möglich (vgl. [Gri75]). In dynamischen Situationen sind auch die Ressourcen und hierbei besonders die Zeit zu berücksichtigen, die dem Sprecher für die Beschreibung zur Verfügung steht. Dadurch hat der Sprecher die Aufgabe, eine Beschreibung zu generieren und den gesamten Vorgang mit der Situation, in der er sich befindet, zu koordinieren.

Natürlichsprachliche Beschreibungen sind inhärent sequentiell, was erfordert, daß der Sprecher festlegen muß, in welcher Reihenfolge er die Nachricht mitteilen will. Dieses wird allgemein als

⁶Die Steuerung muß nicht in jedem Falle sprachlich sein. Im Alltag werden häufig Gesten verwendet, um die Aufmerksamkeit des Adressaten auf Bereiche im deiktischen Raum zu lenken.
Linearisierungsproblem bezeichnet (vgl. [Lev89, S. 138]). Verschiedenen Diskursarten, wie beispielsweise Wegbeschreibungen, sind Prozesse assoziiert, über welche sich eine lineare Ordnung etablieren läßt. Das klassische Beispiel hierzu sind Wohnungs- (vgl. [LL75], [UE82]) und Wegbeschreibungen (vgl. [Kle89], [Wun78]), bei denen die Beschreiber das Konzept der imaginären Wanderungen verwendeten, um eine Beschreibung zu linearisieren. Ein ähnlicher sequenzierendes Konzept läßt sich in noch deutlicherer Form bei inkrementalen Wanderungen nachweisen (vgl. [Ma94]), in denen die Bewegung des Sprechers nicht nur die Reihenfolge, sondern auch die Zeitalstände zwischen Beschreibungen beeinflußt.


Die bisher skizzierten Theorien zur visuellen Wahrnehmung, zum Gedächtnis und zur Sprachverarbeitung bilden einen groben Rahmen dessen, was ein Sprecher leistet, um raum- und sprachliche Beschreibungen zu generieren. Raumkognitive Funktionen basieren auf Leistungen, die durch diese ansatzweise formalisiert sind. Bisher gibt es keine Theorie der Raumkognition, die auch nur annähernd all relevanten Elemente integriert. Im nachfolgenden Kapitel werden die Elemente diskutiert, die notwendigerweise in einer Theorie der Raumkognition enthalten sein müssen. Dabei wird ein besonderer Fokus auf den Aufbau räumlichen Wissens aus visuellen Daten und der Verwendung dieses Wissens für raum- und sprachliche Beschreibungen, insbesondere inkrementale Wegbeschreibungen, gelegt.

Zum einen lassen sich Prozesse identifizieren, die das visuelle System betreffen. Die Schritte, welche notwendig sind, um ein Objekt zu erkennen, die Unterscheidung in Hintergrund und Vordergrund, der Aufbau von Referenzsystemen, anhand derer räumliche Lagebeziehungen zwischen Objekten und dem Wahrnehmenden und auch zwischen Objekten bestimmt werden können. Dies reflektiert eine physikalistische Betrachtungsweise. Nimmt man eine konzeptuelle Ebene zwischen Ein- und Ausgabenmodulen an (siehe Abbildung 2.2), so stellt sich die Frage, wie die Repräsentationen auf beiden Ebenen einander zugeordnet werden. Miller und Johnson-Laird nehmen dazu eine pragmatische Sichtweise ein:

"One useful way of it is that perceptual judgments anew the result of applying conceptual knowledge to sensory inputs." (vgl. [MJL76, S. 60]).

Betrachtet man das raumkognitive System als abgegrenzten Teil des gesamten kognitiven Systems (z.B. [MJL76], [Bry92]), so gibt es afferente Funktionen, die raumbezogene Information
an das raumkognitive System liefern, und eff erente, die sie weiter verwenden. Innerhalb des raumkognitiven Systems vereinen sich Fähigkeiten, wahrgenommene und imaginierte räumliche Information in räumliches Wissen zu transformieren und zu verarbeiten. Da es aus Effizienz- und Beschränkungsgründen nicht notwendig ist, permanent das gesamte räumliche Wissen aktiviert zu haben, ist es indiziert zugreifbar (z.B. [HJ85]).

Es lassen sich vier verschiedene Grundklassen räumliches Wissens unterscheiden:


2. **Referenzsysteme**: Strukturierende Elemente räumlichen Wissens

3. **Räumliche Relationen**: Integrierende Elemente, die Objekte mittels Referenzsystemen qualitativ oder quantitativ zueinander in Beziehung setzen.

4. **Räumliche Konfigurationen**: lokale und globale räumliche Integrationen von Lokationen, bzw. Objekten mittels räumlichen Relationen bezüglich eines oder mehrerer, miteinander verbundener Referenzsysteme


Grundlage für die Formalisierung raumkognitiven Wissens sind Konzepte der Euklidischen Geometrie, wie Achsen, Flächen, Richtungen, Distanzen und Lokationen im dreidimensionalen Raum. Dies beruht auf der Überlegung, daß die Euklidische Geometrie die wesentlichen Eigenschaften des Raumes, so wie ihn der Mensch erfährt, konzeptualisiert. Es läßt sich darüber spekulieren, ob die Euklidische Geometrie das Grundgerüst raumkognitiver Leistungen modelliert (vgl. [OB83, S. 250], [Jac87]). Aufgrund der intuitiven Plausibilität und der einfachen Formalisierbarkeit ist die Euklidische Geometrie die Grundlage für die meisten formalen Modelle im Kontext der Raumkognition.

Für den Aufbau räumlichen Wissens wird davon ausgegangen, daß eine Transformation von Wissensstrukturen perzeptueller und linguistischer Ebenen auf die der raumkognitiven Ebene stattfindet (z.B. [PI67], [Cas57], [Arm74], [MIL76], [OB83], [Bry92]). Die angenommenen Funktionen zwischen den Ebenen differieren in der Komplexität, wenn als Grundlage ein propositionales Format (vgl. [MIL76], [OB83]) oder einer Mischung aus einem propositionalem und einem depiktionalen Format (vgl. [Kos94]) angenommen wird.

---

7 Topologische Relationen benötigen keine Verankerung in einem Referenzsystem.
2.1. ZUSAMMENHANG ZWISCHEN UMGEBUNGEN UND SITUATIONEN

Es wird angenommen, daß die Verankerung von sprachlichen Raumbeschreibungen in visuellen Daten (siehe Abbildung 2.2) durch Vermittlung eines oder mehrerer raumkognitiver Systeme erfolgt (z.B. [ML76], [OB83]). In Anlehnung an Olson und Bialystok lassen sich auf den drei Ebenen verschiedene Prozesse klassifizieren (vgl. [OB83]):

1. **Perzeptuelle Ebene**: automatische und nur gering steuerbare Prozesse auf der niederen und höheren Ebene visueller Verarbeitung; nicht explizierbare Wissensstrukturen

2. **Raumkognitive Ebene**: artikulierbare und separierbare Einzelpropositionen, die intentional und gewollt gesteuert werden können; explizierbare Wissensstrukturen

3. **Sprachebene**: lexikalisierte räumliche Propositionen, die an den Diskurs angepaßt sind; explizierbare Wissensstrukturen

Perzeptuelle Prozesse sind nur in geringem Ausmaße steuerbar. Wissensstrukturen auf der perzeptuellen Ebene lassen sich nach Olson und Bialystok nicht explizieren (vgl. [OB83]). Perzeptuelle Information wird auf der raumkognitiven Ebene in räumliches Wissen transformiert. Sprachorientierte Funktionen haben auf dieses Wissen in ausreichendem Maße Zugriff.


2.1 Zusammenhang zwischen Umgebungen und Situationen

"The geographic world surrounding us is extremely complex. When we want to master a given problem in this world, we need to single out particular aspects of current interest from this multifaceted formation. So at any given time we are only interested in few objects, and concerning these objects again we are regarding only particular properties and/or relations. The capability of isolating the relevant aspects and relating them to one another, results in a unique intellectual efficiency. This efficiency, however, is necessary for successfully operating in the world". (vgl. [FB94, S. 1]).


Eine Umgebung ist durch starke und schwache Dimensionen bestimmt. Schwache Dimensionen sind solche, entlang welcher man sich mit nahezu konstantem Energieaufwand bewegen, kann. Starke Dimensionen zeichnen sich dadurch aus, daß der Aufwand, sich auf dieser zu bewegen erheblich größer ist. Auffällig ist, daß wir in der alltäglichen Realumgebung keine Schwierigkeiten beim Wechsel von verschiedenen dimensionierten Umgebungen haben. Dies mag aber auch daher

---

8Solche natürlichen Umgebungen werden durch von Menschen geschaffene Umgebungen noch erweitert, was die Klasse dessen, was üblicherweise als Umgebung beschrieben wird, nahezu unendlich anwachsen läßt.
ruhren, daß wir uns fast ausschließlich in zwei-dimensionalen, horizontalen Umgebungen fortbewegen (z.B. [Bo163], [Kru74]). Obwohl die meisten Umgebungen des Alltags drei-dimensionale sind, erfolgt die Fortbewegung maximal auf einer zwei-dimensionalen Ebene. Nur in seltenen Fällen spielt eine dritte Raumdimension eine Rolle. Auch wenn jemand einen Berg bestiegt, so führt ihn sein Weg nur scheinbar in eine dritte Raumdimension, da die Ebene, auf der er sich fortbewegt, nur entsprechend der Steigung des Berges angepaßt ist. Vielmehr ist der Weg eines Bergsteigers nahezu durch eine Dimension bestimmt. Im alltäglichen Leben spielt aus diesem Grunde die dritte Dimension nur eine untergeordnete Rolle. Aus diesen Überlegungen läßt sich schließen, daß eine Umgebung des alltäglichen Lebens fast ausschließlich durch zwei Raumdimensionen und zusätzlich einer Zeitdimension repräsentiert werden kann (vgl. [EM95]).


Im folgenden wird räumliches Wissen fokussiert, welches sich aus einem direkt perzipierten Weltausschnitt ableiten läßt.

2.2 Räumliches Wissen in Situationen


10Steigt man eine Treppe hinauf, um zu einer anderen Etage zu gelangen, verwendet man zur Fortbewegung nur eine Raumdimension, nämlich die Höhendimension.

11Zweidimensionale Karten liefern i.a. eine hinreichend adäquate Repräsentation einer dreidimensionalen Umgebung. Eine wirkliche dritte Raumdimension kommt bei einem Tauchgang oder im Weltraum hinzu.

2.3. RAUMKONZEPT

räumliches Wissen einzelner Situationen miteinander integrieren.13
Eine Situation ist selten statisch, da sich zumindest entweder Objekte relativ zum Betrachter oder der Betrachter sich relativ zum Objekt bewegen. Durch Bewegungen verändert sich für den Betrachter die räumliche Struktur der Situation. Der Betrachter versucht, Situationenfolgen zu antizipieren, um bestimmte Vorhersagen über die Umgebung zu machen und adaptiert entsprechend sein Verhalten. Ändert sich eine Situation in einer unvorgesehenen Weise, so muß der Betrachter sich in möglichst kurzer Zeit darauf einstellen, um so die Adäquatheit seines Verhaltens zu gewährleisten. In Situationen, über die der Betrachter Erfahrung besitzt, ist es ihm typischerweise ein Leichtes, sich auf Änderungen einzustellen (vgl. [Nei76]). In unbekannten Situationen kann es dazu kommen, daß der Mensch nicht in der Lage ist, die Situationenfolge in geeigneter Weise zu antizipieren. Hierbei versucht er zumindest, solche Information herauszufiltern, die für die Verhaltensfindung von Bedeutung ist. Die Adaption erfolgt auf der Basis der Analyse externer Merkmale: Andererseits besitzt ein Betrachter Intentionen, die er in einer Situation in geeignete Verhaltensweisen umzusetzen sucht (vgl. [Bra81], [Bra84]). Unterschiedliche Intentionen haben zur Folge, daß der Betrachter auch unterschiedliche Objekte einer Situation fokussiert und identifiziert. So ändert sich u.a. die Menge der etablierten räumlichen Lagebezüge und darüber die mental etablierten räumlichen Vorstellungen.

Im weiteren werden allgemeine Konzepte der Raumkognition diskutiert. Fokussiert werden solche, die für den Aufbau räumlichen Wissens und dessen Verwendung in raumsprachlichen Beschreibungen von Situationen verwendet werden. Die grundlegenden Konzepte der Raumkognition sind die einer Lokation, eines Objektes, eines Referenzsystems und die der räumlichen Relationen. Eine Integration dieser Konzepte ermöglicht es, von einer Situation eine räumliche Vorstellung in Form räumlicher Konfigurationsbeschreibungen abzuleiten. Im folgenden werden diese Elemente in einem größeren Rahmen eingeordnet, der räumliches Wissen feiner strukturiert.

2.3 Raumkonzepte


13Die Unterscheidung in Situationen und Umgebungen ist vergleichbar mit Talmy’s Unterscheidung in fiktive und fakutive Umgebungen (vgl. [Ta96]).
zeitlichen Existenz. Ein Vorkommen spannt dabei ein Spektrum auf, welches von permanent bis kurzfristig reicht. Vorkommen mit eindeutigen Identitäten, wohlspezifizierten Lokationen, verläßlichen und akzeptierten Größenordnungen sowie ausreichender Permanenz bezüglich des betreffenden Zeitbereichs scheinen die größte Kapazität in der Verankerung räumlicher Wissensstrukturen zu besitzen (vgl. Abbildung 2.3a und [Go91, S. 38],[SW75],[PMS88]).

Abbildung 2.3: Hierarchie allgemeiner räumlicher Konzepte

Abbildung 2.4: Konzepthierarchie einiger räumlicher Konzepte bezüglich städtischer Umgebungen
2.3. RAUMKONZEPTE


Vier grundlegende Raumkonzepte sind mit Vorkommen und Verteilungen assoziiert (vgl. [Go91]):

1. Distanz: räumliche Nähe und räumliche Streuung
2. Verbindung: Kontakt, räumliche Sequenz und räumliche Ordnung
3. Gebiet: Inklusion, Exklusion und räumliche Menge
4. Räumliche Stratifikation: räumliche Hierarchien

Ein vom „fundamentalen Gesetz der Geographie“ abgeleitetes Axiom ist das der räumlichen Separierbarkeit, welches durch das Konzept der Distanz beschrieben wird. Damit verbunden sind die Konzepte der räumlichen Nähe bzw. der räumlichen Streuung. Ersteres ist ein Maß für die Nähe zwischen Vorkommen und Verteilungen und zweiteres ein Maß für die Distanz.\(^\text{15}\)


Ein viertes räumliches Konzept, welches auf sprachlicher Ebene von Bedeutung ist, ist das der räumlichen Stratifikation. Bezüglich allen Entitäten einer Umgebung werden räumliche Hierarchien verwendet, die eine Ordnung erzeugen.\(^\text{16}\) Das Konzept der räumlichen Stratifikation ist auch ein organisatorisches Konzept für das Erlernen und Memorieren räumlichen Wissens (vgl. [HJ85]).

In verschiedenen entwicklungspsychologischen Theorien der Kognition, und insbesondere der Raumkognition wird zwischen verschiedenen Stufen der Aneignung räumlichen Wissens unterschieden (vgl. [SW75, She62, Go178]). Primär ist danach die egozentrische Orientierung, die im Laufe der ontogenetischen Entwicklung zu absoluten Orientierungen erweitert wird. Entsprechend dieser Entwicklungsabfolge lassen sich auch unterschiedliche Wissensstrukturen unterscheiden. Der kognitions-wissenschaftlichen Literatur folgend, die zwischen deklarativen und operativen Wissen unterschiedet, wird das räumliche Wissen in Landmarkenwissen und Routenwissen unterteilt. Zusätzlich deuten die Ergebnisse aus zahlreichen empirischen Studien auf

---

\(^{15}\) Ähnliche Konzepte werden durch Proximität, räumliche Variation, räumliche Ähnlichkeit bzw. Unähnlichkeit, Bereichsbildung, nächste Nachbarn und räumliche Heterogenität ausgedrückt (vgl. [Am88]).

\(^{16}\) Beispielsweise wird in Deutschland zwischen Katen, Ein-, Zwei- und Mehrfamilienhäusern, Villa und Palast unterschieden, die sich auf die Hierarchisierung von Behausungen beziehen.
eine dritte Form des räumlichen Wissens hin, welches als Überblicks- oder Konfigurationswissen bezeichnet wird (vgl. [SW75], [She62]).

Im weiteren wird zuerst darauf eingegangen, wie durch Etablierung verschiedener Referenzsysteme und räumlicher Relationen zwischen Objekten und Lokationen räumliches Wissen aufgebaut und sprachlich verwendet wird. Anschließend werden die drei Wissensarten Landmarken-, Routen- und Überblickswissen diskutiert, die sich von situationsbezogenem räumlichem Wissen ableiten lassen.

2.3.1 Räumliche Referenzsysteme


Weiterhin sind im Kontext sprachlicher Beschreibungen räumlicher Sachverhalte die Verwendungsformen von Referenzsystemen vielfach betrachtet worden (z.B. im Kontext von „Deixis in Phantasma“ (vgl. [B34]), Versetzungseides (vgl. [Sen85]) oder sekundäre Deixis (vgl. [Fil71], [Kle82], [Tal83]).

Basierend auf egozentrischen, intrinsischen und extrinsischen Referenzsystemen bestimmt der Betrachter seine Lokation bezüglich seiner Umgebung. Dies ist der Ausgangspunkt für den Aufbau räumkognitiver Repräsentationen (vgl. [HM73]). Entwicklungspychologische Arbeiten zeigen, daß die Fähigkeit, Ordnungsrelationen bezüglich eines egozentrischen Referenzsystems zu etablieren, bei Kindern laut Piagets Studien im Alter von sechs Jahren zu entstehen beginnt. Erst im Alter von ungefähr acht Jahren erlernen Kinder die Fähigkeit, auch außerhalb des Körpers verankerte, intrinsische und extrinsische Referenzsysteme zu verwenden (vgl. [Pia26], repliziert durch [SB55]).

Eine elementare physikalische Eigenschaft, die für die Struktur eines Referenzsystems konstitutiv ist, ist das Konzept der Gravitation G (oben/unten-Achse, siehe Abbildung 2.5). Durch spezielle Funktionen des räumkognitiven Systems hat es der Mensch erlernt, Objekten, deren Hauptsachsen nicht entlang der Gravitationsrichtung ausgerichtet sind, allozentrische Referenzsysteme zuzuordnen (vgl. [OB83]). Die zweite dominante Achse ist die Links/Rechts-Achse, welche orthogonal zur Oben/Unten-Achse parallel zum Gesichtsfeld bzw. den Projektionsebenen der Retinae verläuft (z.B. [HGH91]). Die dritte Achse wird durch die Frontalrichtung der aktuellen Aktion oder Perzeption festgelegt (Vorne/Hinten-Achse, siehe Abbildung 2.5). Eine solche Strukturierung des Raums definiert Regionen um das Zentrum eines Referenzsystems, weswegen es gebietskonstituierend ist (vgl. [Pri91]). In der horizontalen Ebene werden die Gebiete unterschieden, die vor, hinter, links und rechts, bzw. nördlich, östlich, westlich oder südlich vom Ankerpunkt des Referenzsystems liegen.

Es gibt Evidenzen, daß ein Akteur zu einem Zeitpunkt nicht nur über ein einziges, sondern über eine Hierarchie von Referenzsystemen verfügt, die er nach Bedarf etabliert und miteinander vergleicht (vgl. [MIL76], [OB83], [CRI93]). Referenzsysteme werden nach Levinson danach unterschieden, ob sie veränderlich (relativ) oder konstant (absolut) sind (vgl. [Lev96]). Absolute

Im weiteren wird nur die Klasse der relativen Referenzsysteme betrachtet, da ausschließlich diese im Korpus der inkrementellen Wegbeschreibungen nachgewiesen werden konnten. Relative Referenzsysteme lassen sich in egozentrische und allozentrische unterteilen. Bei allozentrischen Referenzsystemen wird entsprechend ihrer sprachlichen Verwendungsweise zwischen intrinsischen und extrinsischen unterschieden:

1. **Egozentrische Referenzsysteme**: Sie sind im Sprecher verankert (z.B. [HBN87], [Ehr85], [Her86], [Gar89], [Lev84], [RS88], [BTF92]) und werden durch intrinsische Eigenschaften des Sprechers, beispielsweise in Abhängigkeit von der retinalen Projektionsfläche, der Ausrichtung des Kopfes oder Körpers determiniert (vgl. [MJJ76], [CRI93]). Andere Bezeichnungen sind *deiktisch* oder *betrachterzentriert*.

2. **Intrinsisch definierte Referenzsysteme**: Sie sind allozentrische Referenzsysteme, die in einer perzipierten oder imaginierten Lokation, insbesondere einem Objekt oder anderen Personen, verankert sind (vgl. [Fil71], [Lev84], [Her86], [Gar89], [FL90], [Sch93]). Die Ausrichtung eines intrinsischen Referenzsystems wird durch Eigenschaften des Objektes determiniert.

---

3. *Extrinsisch definierte Referenzsysteme*: Sie sind ebenfalls allozentrische Referenzsysteme und werden durch globale Eigenschaften einer Umgebung, wie einem Raum, einer Stadt oder einem Bild, oder durch den Betrachter determiniert (vgl. [OB83], [SH84], [CRI93]).

Egozentrische Referenzsysteme (for_e) sind im Egozentrum eines Betrachters verankert (vgl. [HT66]). Ein egozentrisches Referenzsystem unterteilt den Raum um den Betrachter in Gebiete bzw. im dreidimensionalen Raum in Volumina (siehe Abbildung 2.6 für einen 2D-Schnitt).

![Abbildung 2.6: Durch ein egozentrisches Referenzsystem for_e konstituierte Gebiete in der horizontalen Ebene](image)

Mittels intrinsischen Referenzsystemen werden Objekte zueinander in Beziehung gesetzt, ohne direkt Bezug auf das Egozentrum des Betrachters zu nehmen (z.B. [MJL76], [LJ93], [GM95], [Her90]). Es treten aber auch Fälle auf, in denen intrinsische Referenzsysteme in einer Lokation des imaginierter oder realen Raum verankert werden (z.B. [TT92], [Kle79], [Her86]). Im prototypischen Fall wird die Ausrichtung des Referenzsystems durch die Frontalachse festgelegt. Dies gilt insbesondere dann, wenn auf die Gravitationskraft Bezug genommen werden kann und die Transversalrichtung sich von der Frontalrichtung ableitet. Bei fehlender Gravitationskraft zeigt sich eine besondere Präferierung egozentrischer Referenzsysteme (vgl. [FL90]).

Verfügt das Referenzobjekt über *intrinsische* Eigenschaften, legt dies die Frontalrichtung des intrinsischen Referenzsystems a priori fest (siehe Abbildung 2.7). Es lassen sich vier wesentliche Fälle unterscheiden, die eine intrinsische Zuordnung ermöglichen (vgl. [MJL76], [Her90]):

1. Eine Seite des Objektes enthält *Perceptionsfunktionen*
2. Die Seite, die zuerst *sichtbar* wird, wenn sich das Objekt bewegt (z.B. ein Auto)
3. Die *Standardposition* von Personen, wenn sie das Objekt oder die zugeordnete Objektklasse verwenden
4. *Situationsbedingte Eigenschaften*

Besitzt ein Objekt keine referenzierbare, intrinsische Eigenschaft, wird eine Frontseite von *außen* durch den Betrachter zugeordnet (vgl. [MJL76], [Her90], [CRI93]). Dies ist kulturell abhängig und im Deutschen durch die Seite bestimmt, die der Sprecher sieht. In manchen Kulturen wird der Seite die Eigenschaft Front zu sein zugesprochen, die auf der abgewandten Seite liegt (vgl. [Ped93]). Danach lassen sich Sprachen in solche unterscheiden, die ein Referenzobjekt nach einem *Spiegelbild*- bzw. nach einem *Tandemprinzip* ordnen (vgl. [CJ73], [Her86]).

---

18Der Ausdruck „egozentriert“ wurde von Roelofs eingeführt (vgl. [Roe59]).
2.3. RAUMKONZEPTE

Abbildung 2.7: Durch ein intrinsisch determiniertes Referenzsystem \( for_i \) konstituierte Gebiete des Objektes \( A \) in der horizontalen Ebene. Der schwarze Balken in \( A \) gibt ein frontkonstituierendes Merkmal an.

Die horizontale Ausrichtung eines extrinsischen Referenzsystems wird vor allem durch den Kontext der Situation determiniert. Dies kann sowohl durch die räumliche Lagebeziehung zwischen dem Sprecher und dem Referenzobjekt gegeben sein (vgl. [Hil82], [MJJ67], [Sai84]) als auch durch Eigenschaften der Umgebung, die den Sprecher und zu referenzierende Entitäten enthält (vgl. [CRI93]). In Abbildung 2.8 ist eine Situation illustriert, in der ein Betrachter eine räumliche Konfiguration aus einer schrägen Perspektive perzipiert und im Objekt \( A \) ein extrinsisch determiniertes Referenzsystem etabliert. Dabei ist die Front des Objektes \( A \) dadurch bestimmt, daß sie am nächsten zum Betrachter liegt (vgl. [Wun86], [Lev86]).

Auf der Basis dieser Referenzsysteme werden Umgebungen strukturiert. Je nachdem, welches Referenzsystem ein Sprecher einnimmt, generiert er Raumbeschreibungen durch deiktischen, intrinsischen oder extrinsischen Gebrauch von räumlichen Relationen (vgl. [MJJ67], [RS88][15] und Kapitel 2.5). Arbeiten über Referenzsysteme lassen sich in solche unterteilen, die eine deklarative Zuordnung von räumlichen Konfigurationen zu Referenzsystemen vornehmen (vgl. [MN78], [Fli87], [CRI93], [FL90]) und solche, die die Verwendung von Referenzsystemen als Grundlage raumsprachlicher Beschreibungen untersuchen (vgl. [Fil75], [Cla73], [MJJ67], [Tal83], [Gar89], [FTC92], [L93], [HT95]). Es gibt Evidenzen, daß raumsprachliche Beschreibungen nicht nur jeweils auf einem Referenzsystem, sondern auf mehreren gleichzeitig beruhen. Es werden solche Beschreibungen präferiert, die in möglichst vielen Referenzsystemen verankert sind. Für die englischen Präpositionen “above” und “under” konnte dabei nachgewiesen werden, daß extrinsische Referenzsysteme gegenüber intrinsischen bei der Wahl einer vertikalen Orientierung präferiert werden (vgl. [CRI93]). Nach Carlson-Radvansky und Irwin hat die Verwendung eines egozentrischen Referenzsystems scheinbar nur geringe Auswirkungen auf die Wahl einer Beschreibung. Dies widerspricht in gewisser Weise den Ergebnissen, daß Umgebungen primär egozentrisch

---

[15] In diesen Arbeiten wird nicht deutlich, ob räumliche Relationen nur Kernbedeutungen räumlicher Präpositionen sind oder ob es sich um Repräsentationen von Lagebeziehungen zwischen Objekten handelt. Vor allem bei Ratz-Schmidt wird zwischen verschiedenen Betrachtungsweisen gewechselt ([RS88]). Im ersten Falle ließe sich fragen, wovon räumliche Relationen abgeleitet werden. Dazu müßte eine weitere Klasse von sprachunabhängigen, räumlichen Relationen angenommen werden. Im zweiten Falle lassen sich räumliche Relationen als strukturelle Beziehungen zwischen Entitäten auflassen, so wie es visuelle Wahrnehmungstheorien postulieren (z.B. [Mar82], [Koe90], [MZ92]). Bisher ist diese Frage nach dem Status räumlicher Relationen weitgehend unbeantwortet.
Abbildung 2.8: Gebiete, die durch ein extrinsisch determiniertes Referenzsystem $for_e$ bezüglich Objekt A in der horizontalen Ebene konstituiert werden.

erfahren werden (vgl. [HBN87]). Aufgrund kognitiver Ökonomiekriterien ist alternativ die Interpretation möglich, daß egozentrische Referenzsysteme sowohl in die Bestimmung intrinsicer, als auch in extrinsische Referenzsysteme eingehen. Dadurch ließe sich erklären, daß egozentrische Referenzsysteme eine quasi „neutrale“ Wirkung besitzen.

Ein Problem bei der Untersuchung von Referenzsystemen besteht darin, daß sie i.a. mittels sprachlicher Beschreibungen evaluiert werden. Dabei sind Nebeneffekte zu identifizieren, die rein sprachlich-kultureller Natur sind. Entgegen den Evidenzen, daß es mit wachsendem Rotationswinkel schwieriger wird, sich in ein Objekt oder eine Person hineinzuversetzen (vgl. [Her87], [HGH91]), wird in einem kommunikativen Kontext eine vis-à-vis Situation bevorzugt, in der die Gesprächspartner sich gegenüberstehen (vgl. [Cha73], [Meh65], [Som69]). Offenbar kommt es bei der Verwendung von Referenzsystemen zur Bestimmung raumsprachlicher Beschreibungen zu Wechselwirkungen mit kommunikativen Präferenzen. Ein Sprecher möchte einem Rezipienten vielfach dadurch entgegenkommen, daß er dessen Perspektive, d.h., ein extrinisches Referenzsystem annimmt (vgl. [Her87], [Sch93], [Sch95b]). Es ergibt sich beim Vergleich von Reaktionszeiten, daß in bestimmten Situationen eine intrinsische Beschreibung aus der Perspektive des Hörers länger dauert als aus der Perspektive eines Objektes (vgl. [Sch93], [Sch95b]). Da in diesen Untersuchungen keine explizite Unterscheidung zwischen einer Kommunikations- und einer Lokalisation aufgabe gemacht wird, ist es nicht entscheidbar, ob dieser Effekt stärker kommunikativ-kulturellen Ursprungs ist, oder ob er auf der Basis eines neuen Referenzsystems erklärt werden muß.

In den meisten Studien zur Untersuchung von Referenzsystemen wird auf Effekte fokussiert, die auftreten, wenn ein Referenzsystem bereits ausgewählt worden ist (z.B. [MJL76], [Her86]). Es gibt bisher nur wenige Aufschlüsse über Mechanismen zur Etablierung von Referenzsystemen, zum Wechsel zwischen Referenzsystemen, sowie der Verwaltung von Referenzsystemen (z.B. [HGH91]). Aus linguistischer Sicht ist vielfach untersucht worden, welche Referenzsysteme konstituiert werden und wo diese relativ zur beschriebenen Situation verankert sind (z.B. [MJL76], [BTF92]). Es ergibt sich dabei eine Asymmetrie der Reaktionszeiten für räumliche Präpositionen der Frontalachse\footnote{Die Reaktionszeiten für eine Referenzierung durch “in front of” wird schneller gegeben als eine durch “behind”.} für intrinsiche Beschreibungen (vgl. [BTF92]). Diese Asymmetrie ist im
extrinsischen Fall nicht erkennbar. Beide Referenzsysteme sind in einer imaginären Lokation des Betrachters verankert, weswegen sie im erweiterten Sinne als egozentrisch angenommen werden können.

Zur Formalisierung von Referenzsystemen werden zuzüglich zu Eigenschaften der Referenzobjekte auch Information über die betrachterzentrierte, perspektivische Sicht des Sprechers verwendet, um Objekte in Gebiete zu zerlegen. (vgl. [Pri93, S. 178/9]). In einem anderen Ansatz wird entsprechend der Ausrichtung des egozentrischen Referenzsystems durch die geometrischen Eigenschaften eines Referenzobjektes ein umschließendes Rechteck definiert (vgl. [RS88], [HRA89]). Pribenows Vorschlag ist es, das egozentrische Referenzsystem auf das Referenzobjekt zu induzieren und im egozentrischen Fall die Zuschreibungen der vorderen und hinteren Seite, entsprechend dem Spiegelprinzip (vgl. [RS88]) um 180° horizontal zu drehen (vgl. [Pri93, S. 178/9]). In beiden Fällen ist nicht ersichtlich, wie der Zusammenhang zwischen egozentrischen, intrinsischen und extrinsischen Referenzsystemen hergestellt wird. In diesen Arbeiten wird angenommen, daß die vertikale Ausrichtung a priori durch die Schwerkraft vorgegeben ist.

Bezüglich der Verwendung von Referenzsystemen gibt es Evidenzen, daß ein Sprecher mental durch Rotationen und Translationen allozentrische Referenzsysteme einnehmen bzw. aufbauen kann (vgl. [HBN87], [HGH91]). Es bietet sich zunächst an, Arbeiten zum Problem der mentalen Rotation heranzuziehen (z.B. [Cor82], [FS86], [How86]), in welchen die Fragestellung untersucht wird, wie ein Proband ein Objekt mental derart rotiert, daß es deckungsgleich mit einem Objekt B ist. Im Gegensatz dazu wird bei Lokalisationsaufgaben eine Transformation des Blickpunktes des Sprechers und nicht eines Objektes mental untersucht. In zwei Experimenten konnten Hermann und Kollegen Effekte bei der Lokalisation unter Verwendung eines egozentrischen und eines in einem Objekt verankerten, allozentrischen Referenzsystems nachweisen, die nicht mit den Ergebnissen aus den Shepard-Metzler-Experimenten übereinstimmten (vgl. [HGH91]).

Abbildung 2.9: Versuchsanordnung und einzeichneten Elastizitäts- und vis-à-vis-Bereich (vgl. [HGH91])

Intrinsische Lokalisation: Reaktionszeiten in Abhängigkeit vom Rotationswinkel α (vgl. [HGH91]) [hermann-intrin]

Es wurde getestet, wie schnell Probanden eine räumliche Lagebeziehung mittels „links“ oder „rechts“ verbalisieren konnten (siehe Abbildung 2.9). Der große Kreis symbolisiert einen Tisch.

---

21Hermann et al. nennen dieses Phänomen das „Sichhineinversetzen“ (vgl. [HGH91]).
an dem Personen sitzen und Karten spielen. Zu jedem Zeitpunkt wird jeweils nur einSpieler angezeigt. In einer Hand hält der Spieler seine Karten (Markierung). Die Ursprungsposition ist durch den Kreis unterhalb des Tisches angegeben. Es wurden zwei Versuchsreihen durchgeführt. In der ersten betrachtete der Proband den Tisch aus Position P heraus (deiktische Perspektive), bezüglich der er die Lage der Markierung anzugeben hatte (α = 0°) .22 In der zweiten Versuchsreihe wurde der Proband aufgefordert, die Position des Spielers einzunehmen (intrinsische Perspektive).23 Im deiktischen Fall wird das Referenzsystem24 durch die egozentrische Frontalrichtung bezüglich Position P definiert; im zweiten Fall wird die intrinsische Frontalrichtung25 bezüglich der Lokation des Spielers verwendet.

Das Ergebnis der deiktischen und intrinsischen Lokalisation zeigt, daß bei einem Winkel α zwischen 0° und 60° die Reaktionszeiten nahezu konstant bleiben (siehe Abbildungen ?, 2.10). Ist α größer als 90°, so steigt die Reaktionszeit im intrinsischen Fall stark an und erreicht ihr Maximum bei 180°. Im deiktischen Fall erreichen die Reaktionszeiten ein Maximum um 120° (vgl. Abbildung 2.10). Zwischen 120° und 180° kommt es wieder zu einem Rückgang auf ähnliche Reaktionszeiten, wie im Bereich zwischen 0° bis 60°. Dies unterstützt die Hypothese, daß sich die Probanden bei intrinsischer Lokalisation bei kleinem α (zwischen 0° und 60°) ohne großen kognitiven Aufwand in den Spieler hinzuversetzen konnten (Elastizitätsbereich). Schließen die

---

22In Abbildung 2.9 ist die Markierung bei einer deiktischen Beschreibung auf der rechten Seite.
23In Abbildung 2.9 ist die Markierung bei einer intrinsischen Beschreibung auf der linken Seite.
24In Abbildung 2.9 sind die Referenzsysteme in der horizontalen Ebene durch doppelläufige Pfeile eingezeichnet. Das egozentrische Referenzsystem ist in der Lokation des Probanden und das intrinsische in der des Spielers verankert.
25In Richtung auf den Mittelpunkt des Tisches
2.3. RAUMKONZEPTE

Abbildung 2.10: Deiktische Lokalisation: Reaktionszeiten in Abhängigkeit vom Rotationswinkel α (vgl. [HGH91])

deiktische und die intrinsische Frontalrichtung einen größeren Winkel ein, so sind im intrinsischen Fall winkelproportional aufwendigere kognitive Leistungen vonnöten, um die Markierung zu lokalisieren. Im extrinsischen Fall wird um 120° ein Reaktionszeitenmaximum erreicht, worauf Herrmann et al. auf aufwendigere kognitive Leistungen für die Zuordnung des perzipierten Tokens zu einer der Markierungskategorien (links oder rechts) schließen. Im vis-à-vis-Bereich kommt es im deiktischen Fall zu einer Verringerung der Reaktionszeiten, da die Lokalisationsaufgabe ähnlich zu der im Elastizitätsbereich ist. Die Ergebnisse dieser beiden Versuche zur Selbstrotation stehen im Gegensatz zu der quasi-linearen Beziehung zwischen Rotationswinkel und Reaktionszeiten in den Untersuchungen mentaler Rotationen von Objekten. Aus dieser Experimentalserie läßt sich ableiten, daß die besten Leistungen zur Objektlokalisation, sowohl im deiktischen, als auch im intrinsischen Falle, im Elastizitätsbereich von 0° bis 60° möglich sind. Diese Ergebnisse für intrinsische Beschreibungen sind von Schober repliziert worden (vgl. [Sch93], [Sch95b]).

Die meisten der genannten Arbeiten fokussieren die Verwendung von Referenzsystemen in statischen Situationen. Jedoch sind es meist dynamische Situationen, mit denen ein Sprecher im alltäglichen Leben konfrontiert wird. Durch Bewegungen werden Erwartungen für zukünftige Situationen vorgegeben. Eine Extrapolation der Trajektorie eines Objektes erlaubt die Antizipation einer Bewegung, so daß Situationen vorhergesagt werden können. Bewegt sich ein Sprecher real durch eine Umgebung, so nimmt er primär ein egozentrisches Referenzsystem als Grundlage seiner Verständnisse von Raum an (vgl. [PM94], [Rie89]).

2.3.2 Räumliche Relationen

Die Lokation eines Objektes wird für einen Betrachter erst dann verständlich, wenn sie durch räumliche Relationen zu einem perzipierten, memorierten oder bereits imaginingen Referenzobjekt in Bezug gesetzt worden ist. Räumliche Relationen lassen sich formal durch ein dreistelliges Prädikat explizieren ($R(\text{ego}, \text{lo}, \text{relo})$). Der Term ego repräsentiert die Lokation des Betrachters, wodurch ein direkter Bezug zwischen der Relation $R$ und dem Betrachter hergestellt wird. In den meisten Arbeiten zu räumlichen Relationen wird dieser direkte Bezug zum Sprecher nur implizit betrachtet (z.B. [MJL76], [Her86]). Unter der Annahme, dass intrinsische und extrinsische Referenzsysteme von egozentrischen Referenzsystemen abgeleitet werden, ist es notwendig, die Lokation des Betrachters in der Formalisierung der räumlichen Relationen direkt zu berücksichtigen.\textsuperscript{26} Der Betrachter ist der Ursprung für die Etablierung einer räumlichen Relation.\textsuperscript{27} Der zweite Term lo steht für das Lokalisationsobjekt. Der dritte Term relo expliziert das Referenzobjekt, bezüglich dem das Lokalisationsobjekt in Beziehung gesetzt wird.\textsuperscript{28} Eine solche Figur-Grund-Eigenschaft von räumlichen Relationen zeigt eine gewisse Nähe zu Eigenschaften bei der Separierung von Objekten im Kontext der visuellen Wahrnehmung.

Nach Pullar und Egenhofer [PE88] lassen sich räumliche Relationen in vier Klassen unterteilen:

1. Topologische Relationen: Lagebeziehungen zwischen zwei Objekten, die Disjunktheit, Kontakt, Überlappung oder Inklusion explizieren; topologische Relationen benötigen kein Referenzsystem (für Formalisierungen vgl. [Ege91], [EAT92], [Coh87], [RCC92], [Rup96])

2. Deiktische Relationen: repräsentieren Lagebeziehungen zwischen dem Sprecher und Objekten (vgl. [MJL76], [RS88])

3. Ordnungsrelationen: Lagebeziehungen, die Objekte anhand einer Ordnung sequentialisieren

4. Metrische Relationen: explizieren den metrischen Aspekt einer Lagebeziehung, wobei die verwendete Metrik vom Sprecher und Hörer in gleicher Weise verstanden sein muß (z.B. [McN92])

Eine andere Unterscheidung ist es, räumliche Relationen in quantitative und qualitative räumliche Relationen einzuteilen (vgl. [FR93]). Metrische Relationen sind rein quantitativ, wohingegen die ersten drei Klassen räumlicher Relationen qualitativ sind. Quantitative räumliche Relationen sind lokationskonstituierend, wohingegen qualitative räumliche Relationen gebietskonstituierend sind. Qualitative Relationen geben keine genaue Lokation, sondern ein Gebiet an, welches prototypisch dieser Relation in einer gegebenen Situation zugeordnet wird. Ein Beispiel für eine quantitative Relation ist eine Relation $R_1$, die repräsentiert, daß zwei Lokationen $x$ und $y$ sich genau in einem Abstand von einem Meter befinden. Ein Beispiel für eine qualitative Relation ist eine solche, die auf sprachlicher Ebene durch die Präposition „nahe“ ausgedrückt wird.

\textsuperscript{26}Ein weiterer Vorteil dieses Ansatzes ist es, daß räumliche Relationen in ihrer Stelligkeit nicht variieren, sondern homogen behandelt werden können.

\textsuperscript{27}Bei Sekundärdeixen wird die Lokation des Sprechers ego durch eine Referenz zu einer Lokation in dieser lokalen Umgebung ersetzt (siehe [Kle70]). In der Beschreibung „Schaust man von Berkeley aus nach San Francisco, so liegt Palo Alto in einer Relation L von San Francisco,“ nimmt der Sprecher eine imaginäre Position ein, die in diesem Fall ganz Berkeley umfaßt. Es gilt somit die Relation $L(\text{loc(Berkeley)}, \text{loc(Palo Alto)}, \text{loc(San Francisco)})$. Zwischen der Lokation, die durch loc(Berkeley) repräsentiert wird, und dem Sprecher gibt es dadurch einen kausalen Zusammenhang, daß sich der Sprecher mental in die Lokation von Berkeley versetzt. Es gilt somit zwischen der Lokation des Sprechers und Berkeley zusätzlich eine Transitionsrelation $S(\text{ego}, \text{loc(Berkeley)})$. Dadurch gilt die Relation L dann, wenn der Sprecher seine Lokation ego imaginär nach Berkeley verlagert. In diesem Fall wird L durch eine links-von Lagebeziehung repräsentiert.

\textsuperscript{28}In der Figur-Hintergrund-Terminologie wird lo als Figur und relo als Hintergrund bezeichnet (vgl. [Ta83]).
Für jede Klasse von räumlichen Relationen werden unterschiedliche Anforderungen an das verwendete Referenzsystem gestellt (siehe Tabelle 2.2). Topologische Relationen benötigen kein Referenzsystem, da die Ausrichtung und Lokation vom Betrachter unabhängig ist (vgl. [MJJL76], [Ege91]), wohingegen deiktische Relationen auf qualitative Eigenschaften eines egozentrischen Referenzsystems zurückgreifen (vgl. [MJJL76], [RS88]). Ordnungsrelationen und metrische Relationen können bezüglich allen Referenzsystemen etabliert werden. Dabei nehmen Ordnungsrelationen eine lokale oder globale Ordnungseigenschaft bezüglich der Dimensionen eines Referenzsystems an. Metrische Relationen setzen voraus, daß durch das Referenzsystem Metriken bezüglich den Dimensionen definiert sind. Eine Eigenschaft einer solchen Metrik ist es, Distanzen zwischen Lokationen entsprechend einer Funktion vorzugeben, was für einen Betrachter ohne Meßinstrumente oder ausreichende Erfahrung i.a. schwierig herzustellen ist. Eine Metrik ordnet jedem Paar zweier Lokationen \( x \) und \( y \) eine Distanz \( d(x, y) \) zu, welche drei Bedingungen erfüllt (vgl. [McN92]):
1. Positivität: $d(x, x) = 0$ und $d(x, y) > 0$, wenn $x \neq y$
2. Symmetrie: $d(x, y) = d(y, x)$
3. Dreiecksungleichung: $d(x, z) \leq d(x, y) + d(y, z)$
4. Additivität der Segmente: $\forall$ Lokationen $x, y, z$: $x$ und $z$ sind durch ein Segment verbunden:
   $\forall$ y auf dem Pfad zwischen $x$ und $z$: $d(x, y) = d(x, y) + d(y, z)$

Metrische Relationen werden im räumlichen Wissen typischerweise verzerrt repräsentiert. Es kommt u.a. zu Asymmetrien der Distanzen zwischen Objekten (vgl. [Ste69], [Cad79], [SBS80]) und zur Verletzung der Additivitätsbedingung.

### Tabelle 2.2: Beziehung zwischen räumlichen Relationen und Referenzsystemen

<table>
<thead>
<tr>
<th>Topologische Relationen</th>
<th>Egozentrisches Referenzsystem</th>
<th>Allozentrisches Referenzsystem</th>
<th>ohne Referenzsystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deiktische Relationen</td>
<td>●</td>
<td>●</td>
<td>•</td>
</tr>
<tr>
<td>Ordnungsrelationen</td>
<td>●</td>
<td>●</td>
<td>•</td>
</tr>
<tr>
<td>Metrische Relationen</td>
<td>●</td>
<td>●</td>
<td>•</td>
</tr>
</tbody>
</table>

Deiktische Relationen werden wesentlich durch egozentrische Referenzsysteme konstituiert. Es werden durch diese Relationen in einer dreidimensionalen Umgebung sechs Gebiete identifiziert: VOR, HINTER, ÜBER, UNTER, RECHTS-VON, LINKS-VON (siehe Abbildung 2.6). Assoziiert mit diesen Gebieten ist eine beschränkte Menge von deiktischen Relationen: #in-front-of#, #behind#, #above#, #under#, #right-of#, #left-of#. Dadurch läßt sich angeben, welche Bedingung erfüllt sein muß, damit eine deiktische Relation $DR$ zwischen einer imaginierter oder realen Lokation des Sprechers $S$ und einem Objekt $O$ erfüllt werden kann (vgl. Definition DEIKTISCHE RELATIONEN).

### Definition 1: Deiktische Relationen

Eine deiktische Relation $DR(loc(S), loc(O), loc(S))$ ist erfüllbar, gdw. $O$ im zu $DR$ assoziierten Gebiet bezüglich eines egozentrischen Referenzsystems liegt.

Ordnungsrelationen lassen sich in ähnlicher Weise über Gebiete bezüglich eines Referenzsystems etablieren (siehe Abbildung 2.7, 2.8 und Definition ORDNUNGSPERATIONEN).

Bei metrischen Relationen reduziert sich die Erfüllbarkeit auf Lokationen (vgl. Definition METRISCHE RELATIONEN). Da sich Gebiete i.a. nicht eindeutig voneinander abgrenzen lassen, werden in manchen Arbeiten sogenannte Anwendbarkeitsgrade eingeführt, die angeben, inwieweit die Lokation eines Objektes in einer prototypischen Region des Gebietes liegt (vgl. [Her86]).

---

29 Diese Bedingung wird eingeführt, um triviale Metrien, wie z.B. $d(x, x) = 0$ und $d(x, y) = 1$, auszuschließen. Eine Distanzmessung, die alle vier Bedingungen erfüllt, wird als "Metrik mit additiven Segmenten" bezeichnet.
30 Die Restriktion der räumlichen Relationen entspricht den Prinzipien der minimalen Repräsentation und dem geringsten Zugeständnis. Eine solche Einschränkung wird auch in der psychologischen Literatur vorgenommen (z.B. [GW95], [Log95], [HKB92]).
2.3. RAUMKONZEPTE

Definition 2: Ordnungsrelationen

Eine Ordnungsrelation OR(loc(S), loc(O1), loc(O2)) ist erfüllbar, gdw. O1 im zu OR
assozierten Gebiet bezüglich eines in O2 verankerten Referenzsystems liegt.

Definition 3: Metrische Relationen

Eine metrische Relation MR(loc(S), loc(O1), loc(O2)) ist erfüllbar, gdw. O1 auf der zu
MR assozierten Lokation bezüglich der Metrik M eines in O2 verankerten Referenzsystems
liegt.

[ABHR87], [Gap94], [Her96]). Ein solcher Ansatz erweist sich besonders bei der Umsetzung auf
eine sprachliche Ebene als vorteilhaft (z.B. [ABHR87], [Her96]).

Mit Objekten, räumlichen Referenzsystemen und räumlichen Relationen sind die Konzepte eingeführt worden, die für den Aufbau räumlichen Wissens notwendig sind. Im folgenden wird
diskutiert, wie im Kontext von Raumbeschreibungen geeignete Referenzsysteme ausgewählt (Abschnitt 2.3.3) und verschiedene Arten räumlichen Wissens aufgebaut werden (2.3.4, 2.3.5, 2.3.6).

2.3.3 Wahl eines Referenzsystems im Kontext von Raumbeschreibungen

Wissen über Lokationen von Objekten in einer Umgebung ist eine notwendige Voraussetzung für
adäquates sprachliches Verhalten. Um auf eine Lokation Bezug nehmen zu können, wird diese
Lokation zu der bereits etablierten räumlichen Struktur in Beziehung gesetzt.

Drei Bedingungen sind für eine erfolgreiche Kommunikation räumlicher Relationen notwendig
(vgl. [Kle94]).

1. Der Sprecher und der Rezipient teilen sich in einem ausreichend hohen Maße einen kognitiv-
rekonstruierten Raum

2. Der Sprecher und der Hörer beherrschen die lexikalischen Bedeutungen raumbezogener
Auszüge

3. Die Kernbedeutung wird mit allen Sorten von Kontextinformation passend unterstützt
(z.B. durch Positionen und Blickrichtungen)

Unter der Annahme, daß der zweite Punkt eine Grundvoraussetzung ist, wird jetzt der erste
Punkt eingehender betrachtet. Inwieweit Kontextinformation verwendet werden kann, wird in
Kapitel 2.5.2 diskutiert.

Eine Voraussetzung für die Verankerung räumlichen Wissens ist die Auswahl geeigneter Referen-
zenysteme. Um einen räumlichen Sachverhalt kommunizieren zu können, muß der Rezipient
Referenzsysteme annehmen, welche denen des Sprechers hinreichend ähnlich sind. Die Wahl
eines geeignetes Referenzsystems ist für den Nutzen einer Beschreibung entscheidend. In der
sprachorientierten Literatur finden sich u.a. sechs Faktoren, die die Wahl eines Referenzsystems
beeinflussen:
1. Verbale Explikation (vgl. [Sai84], [MJL76])

2. Intrinsische Präferenz (vgl. [MJL76], [Abk82], [Cox81], [Lev82], [Wun82])

3. Individuelle Unterschiede (vgl. [Lev82])

4. Konstante intrinsische oder deiktische Beschreibung in einem Diskurs (vgl. [LL75], [vC93], [Ehr85])

5. Statischer oder dynamischer Charakter einer Situation (vgl. [Wun81])

6. Integrierte Betrachtung von situationsbedingten Faktoren und intrinsischen Eigenschaften des Referenzobjektes (vgl. [GM95])

Im ersten Fall wird die Art des Referenzsystems dem Hörer explizit mitgeteilt, so daß die Wahl willkürlich ist. Vielfach untersucht worden sind im Hinblick auf die Wahl eines Referenzsystems, in welcher Weise intrinsische Eigenschaften dies beeinflussen. Obwohl sie einen starken Einfluß haben, kann es dennoch zu Ambiguitäten kommen (vgl. [Her90]), die u.a. kulturabhängig sind (vgl. [Car93], [Kie94]). Die Kategorisierung in Personengruppen, die deiktische, und solche, die intrinsische und extrinsische Beschreibungen bevorzugen, erscheint fragwürdig (vgl. [GM95]). Weiterhin lassen sich Kulturabhängigkeiten nachweisen (z.B. in [Hav92], [Lev92], [Ped93]). Im gleichen Sinne ist eine diskursbedingte, konsistente Verwendung einer Klasse von Referenzsystemen nur in Spezialfällen, wie Wohnungsbeschreibungen, belegbar (vgl. [LL75]). Die Zuordnung von Referenzsystemen entsprechend dem Charakter von Situationen läßt sich nur vage aufrecht erhalten. Vielversprechender ist der letzte Faktor: Grabowski und Weiβ konnten zeigen, daß in einer institutionalisierten Situation in unserer Kultur allozentrische Referenzsysteme bevorzugt werden (vgl. [GW95, S. 12]). In informellen Situationen werden hingegen meist egozentrische Referenzsysteme verwendet. Allgemein wird in der deutschen Sprache in den meisten Fällen eine deiktische Beschreibung gewählt, um ein Referenzobjekt zu indizieren, welches keine intrinsischen Eigenschaften besitzt (vgl. [GM95]). Besitzt das Referenzobjekt intrinsische Eigenschaften, so beeinflußt der situative Kontext und die intrinsischen Merkmale des Objektes gemeinsam, ob ein deiktisches, intrinsicsches oder extrinsisches Referenzsystem verwendet wird (vgl. [GM95]). Andere Arbeiten gehen davon aus, daß die Etablierung eines intrinsischen oder extrinsischen Referenzsystems mit erheblichen kognitiven Aufwand verbunden ist und daβ aus diesem Grund egozentrische Referenzsysteme präferiert werden (z.B. [BNH86], [Ehr85], [HBN87], [HiL82], [Lev82], [Lev84, Wun81], [CRI93]). Interessant ist, daβ einfache raumbezogene Beschreibungen vom Rezipienten zu einem großen Anteil nicht korrekt verstanden werden.31

Auf der Basis von geeigneten Referenzsystemen verbindet ein Betrachter Objekte einer Umgebung durch Etablierung von räumlichen Relationen zu räumlichen Wissenstrukturen. Es werden dabei drei räumliche Wissenstrukturen unterschieden, die sich auf das Wissen über Landmarken, Routen und ganzen Konfigurationen beziehen. In den nächsten Abschnitten werden diese nacheinander diskutiert.

2.3.4 Landmarkenwissen

In jeder Umgebung gibt es Objekte, die in räumlichen Beschreibungen als markant oder auffällig beschriftet werden. Durch Verwendung solcher Objekte strukturiert ein Betrachter seine Umgebung. Solche markanten Objekte werden als Landmarken bezeichnet. Lynch war der erste, der

---

31In empirischen Untersuchungen konnte von Grabowski und Miller gezeigt werden, daß bezüglich einer Einparksituationsschreibung nur 63% aller raumbezogenen Beschreibungen zu einer erfolgreichen Kommunikation führten, in der sich der Hörer und der Sprecher verstehen (vgl. [GM95]). Für die deutsche Sprache zeigt sich, daß Hörer und Sprecher über das Gebiet zwischen beiden und dem Referenzobjekt erheblich besser kommunizieren als über den Raum hinter dem Referenzobjekt (vgl. [GM93]).
2.3. RAUMKONZEPTE

Eine Landmarke als eine raumstrukturenden Einheit auffaßte (vgl. [Lyn60]). Hiernach ist eine Landmarke ein Referenzpunkt, zu dem andere Objekte einer Umgebung in Beziehung gesetzt werden, wodurch sie der Wegfindung in komplexen Umgebungen dienen (vgl. [App70], [BW76], [CS69], [Lyn60], [ASR78]). Eine Voraussetzung ist die räumliche Prommenz eines Objektes, das dadurch entsteht, daß das Objekt von mehreren Seiten sichtbar ist oder sich von anderen Objekten der Situation bezüglich ihrer Merkmale abhebt (vgl. [Lyn60]). Subjektive Einstellungen des Betrachters modifizieren die Prommenz eines Objektes. Liegt ein Objekt beispielsweise in der Nähe einer Kreuzung, an welcher der Betrachter abbiegen möchte, steigert dies die Prommenz des Objektes (vgl. [Lyn60]). In perzipierten Umgebungen ist die Prommenz wesentlich durch visuelle Eigenschaften eines Objektes determiniert (vgl. [PM88]). Die Zuordnung von referentiellen Eigenschaften erlaubt es, andere Lokationen relational mit Landmarken in Beziehung zu setzen (vgl. [SBS80]).

In der psychologischen Literatur wird zwischen Landmarken unterschieden, die in einer Situation direkt perzipiert werden (z.B. [Gib79]) und solchen, die strukturierende Bedeutung in räumlichen Langzeitrepräsentationen besitzen (vgl. [Pre82], [Pre87], [PM88]). Im prototypischen Fall wird eine Landmarke als ein stationäres Objekt klassifiziert, welches eine Lokation im Raum bestimmt (z.B. [App69], [CS69], [SW75], [PM888]). Andererseits können auch bewegliche Objekte wie die Sonne32 als Landmarken fungieren. Entwicklungspyschologisch lassen sich zwei Sichtweisen auf Landmarken identifizieren (vgl. [HM73], [SW75]). Landmarken dienen der Orientierung, der Wegfindung und dem Verfolgen eines Weges (vgl. [Acr88]).

Das Piagetische Konzept einer topologischen Repräsentation ist in diesem Kontext so interpretiert worden, daß Kinder Lokationen nur unter Verwendung von Landmarken, und dabei insbesondere räumlich nahe Landmarken, referenzieren können (vgl. [APO75], [HST78]). Da nur dann eine Aussage über die räumliche Nähe zwischen zwei Objekten gemacht werden kann, wenn vorher der Abstand zwischen diesen beiden bestimmt worden ist, tritt das Paradoxon auf, daß räumliche Nähe, bzw. Nachbarschaft, inhärent mit Distanzen und darüber mit Metrik verbunden sind. Diese offene Frage kann dadurch umgangen werden, daß angenommen wird, daß Kinder, im Widerspruch zu Piaget, über ein metrisches Konzept der Distanz (vgl. [LGS81]), bzw. ein nicht-metrisches Konzept der Nachbarschaft oder eine Fähigkeit der Ordnungsbildung verfügen (vgl. [New88]). Die Bedeutung von Landmarken für die räumliche Orientierungsfähigkeit wird durch neurophysiologische Studien unterstützt. Durch Läsionen des ventralen Systems33 bedingte Probleme, Objekte zu erkennen, haben auch Auswirkungen auf das Erkennen von bekannten Landmarken (vgl. [LWF89]) und dadurch auf die Fähigkeit, Wege zu finden (vgl. [She92]).

Aus Gründen der Ökonomie erscheint es wahrscheinlich, daß nicht alle potentiellen Relationen zwischen Landmarken automatisch etabliert werden (vgl. [Sad88]). Ein effizienter Ansatz ist es, Relationen nur zwischen ausgewählten Referenzpunkten und Landmarken zu bestimmen, wodurch räumliches Wissen hierarchisiert wird (z.B. [SC78], [Tev81], [HJS85]). Dies ist konsistent mit dem Vorschlag, einer Landmarke ein minimales Referenzsystem zuzuordnen (vgl. [PY97]), wodurch sich Lokationen und Landmarken zu Konfigurationen zusammenfassen lassen (z.B. [PM88]). Durch Zuordnung eines Referenzsystems werden Landmarken zur Orientierung und Lokalisierung verwendet (vgl. [PM88], [Nad88]). Eine solche Sicht basiert auf der Annahme, daß raumkognitive Fähigkeiten inhärent relational sind (vgl. [PM88]). Landmarken bewirken typischerweise eine Verzerrung des memorierten räumlichen Wissens. So werden Distanzen wzi-

32In dem Sinne, daß die meisten, möglicherweise nahezu alle, Menschen, entgegen des physikalischen Wissens, die Sonne immer noch als sich um die Erde drehend erfahren.
33Es gibt die Vermutung, daß das visuelle System aus zwei getrennten Systemen besteht, welche in Area V1 des visuellen Kortex verankert sind. Das ventrale System endet im Temporallappen und verarbeitet in wesentlichen Teilen Information, die sich auf die Eigenschaften von visuell wahrgenommenen Objekten bezieht. Das dorsale System hingegen endet neurophysiologisch im Partialappen und verarbeitet hauptsächlich räumliche Information (vgl. [MUM83]). Plakativ werden diese beiden als die „What“- und „Where“-Systeme bezeichnet.
schen Landmarken und Nicht-Landmarken als größer angenommen als umgekehrt (vgl. [Cad79], [HM82], [SBS80]).


2.3.5 Routenwissen


Routenwissen wird häufig als linearer Modus räumlichen Wissens bezeichnet (vgl. [All81b], [HLP78], [PIS60]), da es eine Sequenz von geordneten, zeitlich-räumlichen Beziehungen zwischen einem Akteur und seiner Umgebung expliziert (vgl. [AS78], [BW76], [MC82], [AK85]). Aus Gründen der kognitiven Ökonomie wird postuliert, den kontinuierlichen Fluß zeitlich-räumlicher Erfahrung während des Erlermens einer Route in Segmente (chunks) aufzuteilen (vgl. [AK85], [All81a], [All87]). Das Segmentieren von Routen ist konzeptionell eine Kategorisierung von symbolischen räumlichen Ordnungen. Die Einteilung in Segmente hat Auswirkungen auf das Schätzen von Entfernungen. Empirische Studien zeigen einen positiven Zusammenhang zwischen
2.3. RAUMKONZEPTE

geschätzten und tatsächlichen Entfernungen innerhalb von Segmenten, sowie schwach positive bzw. keine Zusammenhänge zwischen Segmenten (vgl. [AK85]). Mit Routenwissen sind weiterhin Fähigkeiten verbunden, einen Reiseplan vorab auszuwählen, in eine propositionale Struktur zu bringen und Heuristiken zu bestimmen, mit denen dieser Plan in Aktionen umgesetzt werden kann (vgl. [GBL85], [MC82]).


Ein Minimum an Routenwissen repräsentiert eine sequentielle Abfolge von Lokationen und Distanzen zwischen Lokationen, die ein Akteur von der Startlokation bis zur Ziellokation perzipiert, sowie einer Repräsentation der Aktionen an Verzweigungspunkten (vgl. [Kui78], [Lei87], [SL88]). Erweitertes Routenwissen enthält Information über sekundäre und tertiäre Landmarken entlang des Pfades, Distanzen zwischen Landmarken, Relationen zwischen Landmarken und Pfadabschnitten zu einem globalen Referenzsystem (vgl. [Gol91]).

Bei der Transformation von Pfadabschnittsrepräsentationen in räumlichen Langzeitrepräsentationen kommt es zu charakteristischen Verzerrungen und Fehlern (z.B. [Tve81], [Tve92]). Die Segmentierung eines Weges wird um so effizienter durchgeführt, je differenzierter die betreffende Umgebung ist (vgl. [All81]), wobei die Eigenbewegung einen starken Einfluß auf das Erlernen räumlicher Eigenschaften besitzt (vgl. [Acr88], [LPN81]).

2.3.6 Konfigurationswissen

Die dritte Art räumlichen Wissens ist Konfigurationswissen (vgl. [SW75]). Dieses entsteht, wenn Distanzinformation präzisiert wird, sowie Vorstellungen von Winkelzusammenhängen, Richtungen, Kontinuitäten und übergreifende Relationen abgeleitet werden. Dies erfolgt durch Integration von Routensegmenten zu höheren Repräsentationen räumlichen Wissens (vgl. [HMP76], [LSG84], [SW75]).

Konfigurationswissen kann flexibler als Landmarken- und Routenwissen verwendet werden, da Relationen zwischen Landmarken direkt zugreifbar sind. Abkürzungen und Alternativpfade können dadurch einfach bestimmt werden (vgl. [THR82], [LG83]). Insbesondere für Konfigurationswissen gibt es Evidenzen, daß der Zugriff bzw. die Repräsentation hierarchisiert erfolgt (vgl. [HJ85], [McN86], [Sho87]). Konfigurationswissen soll hier nicht eingehender diskutiert werden, da es Bestandteil des Langzeitgedächtnisses ist, welches in dieser Arbeit nicht berücksichtigt wird.


2.3.7 Theorien zu Kognitiven Karten

Für die Gesamtheit individueller, räumlichen Wissens im Langzeitgedächtnis werden verschiedene Begriffe verwendet: "imagery map" (vgl. [Trol3]), "Feldkarte" oder "Kognitive Karte" (vgl. [Toh88], [BBM70], [DS73]), "mentale Karte" (vgl. [Hal55]), "Schema" und "topographische Repräsentation" (vgl. [She62]). Der Begriff der Kognitiven Karte hat sich weit hin durchgesetzt. Das Konzept der Kognitiven Karte expliziert raumbezogene Langzeitrepräsentationen, die als Bindeglied sowohl zwischen Perzeptionen, Umgebungen, als auch raumbezogener sprachlicher Beschreibungen fungieren (vgl. [Tve93], [MHH92b]).
Eine Kognitive Karte integriert Landmarken-, Routen- und Konfigurationswissen, die es einem Akteur erlauben, mentale Vorstellungen und Inspektionen über räumliche Zusammenhänge durchzuführen. Das strukturierende Prinzip der Hierarchisierung in den einzelnen Wissensformaten findet sich auch auf der Ebene der Kognitiven Karten wieder (vgl. [Tve81], [LH87], [Tve92], [MHH92a]), formale Modellierungen finden sich z.B. in [Kos87], [HHD91]). Bei der Verwendung von räumlichem Wissen kommt es zu systematischen Verzerrungen und zu Fragmentierungen der gespeicherten Wissensstrukturen (vgl. [Tve93]). Typische Fehler sind Distanzverzerrungen (z.B. [SBS80], [HM82], [KPF74], [NL82], [SS80], [Tho81]), Gruppierungsfehler (vgl. [Tve81]), Rotationsfehler (vgl. [Tve81], [Cha82], [Llo89], [LH87]) und Linearisierungsfehler (vgl. [MJ76], [SC78], [HM92], [MB83], [SM89]).

Einer der Effekte, die bei der Verwendung räumlichen Langzeitwissens auftreten, sind systematische Fehler bei der Angabe von Richtungen bezüglich Umgebungen (vgl. [SC78]). Landmarken werden räumlichen Gebiete zugeordnet, wodurch Konglomerate und Hierarchien entstehen. Reaktionszeiten bezüglich Richtungsentscheidungs- und anderen räumlichen Aufgaben über Konglomerate hinweg sind kürzer als innerhalb eines Konglomerats (vgl. [Wil75], [Mak81]). Beispielsweise lassen sich bezüglich städtischen und staatenbezogenen Umgebungen Priming-Effekte nachweisen (z.B. [Ste78], [Shu84]). Wird die Schätzung der Distanzen zwischen zwei Lokationen getestet, so zeigen sich systematische Asymmetrien (vgl. [SBS80], [Cad79]). Gleich Distanzen werden zwischen Konglomeraten länger geschätzt als innerhalb eines Konglomerats (vgl. [All81]). Die hierarchische Struktur Kognitiver Karten repräsentiert Referenzzemente, Wissen über Distanzen und semantische Information über Objekte von Umgebungen (vgl. [HJ85], [Tve93]). Eine Hierarchisierung räumlichen Wissens birgt jedoch die Frage in sich, ob eine Repräsentation auf einem zwei- oder dreidimensionalen euklidischen Raum basiert (vgl. [GR73], [GS73], [GRR82], [GZ73], [Rie82]). Empirische Evidenz deutet darauf hin, daß die räumliche Information über wohlbekannte Plätze durch eine zweidimensionale Struktur repräsentiert wird (vgl. [Go87]).

Nachdem die empirischen Befunde im Kontext des räumlichen Wissens skizziert worden sind, werden im weiteren formale Aspekte des motorischen Verhaltens vorgestellt.

2.4 Aspekte des motorischen Verhaltens

Ein weiterer wichtiger Interaktionsmodus mit einer Umgebung öffnet sich durch die Motorik. Aus formaler Sicht kann ein physikalischer Agent über Effekten seine Umgebung manipulieren oder sich darin fortbewegen. Im Hinblick auf sprachliche Beschreibungen einer Umgebung soll hier nur der letzte Fall diskutiert werden.

2.5 Interaktion von Sprache und räumlichem Wissens

Räumliches Wissen repräsentiert das Wissen eines Betrachters über Umgebungen und Situationen. Eine zentrale Frage in der Sprachforschung ist es, in welcher Weise Sprache einerseits dazu verwendet wird, räumliches Wissen aufzubauen (z.B. [JL83], [BT92]) und andererseits in kommunikativen Kontexten zu verwenden (z.B. [Bie67], [Cla73], [Ta83], [Her86]). Im weiteren Verlauf wird ausschließlich Bezug auf den zweiten Aspekt genommen.


Bei der Analyse sprachlicher Beschreibung räumlicher Sachverhalte werden drei grundlegende Konzepte unterschieden: *Entitäten, räumliche Relationen* und *Referenzsysteme*. Als erstes werden *Entitäten* festgelegt, über die eine sprachliche Aussage gemacht wird. Diese können Teil einer größeren Entität oder weitgehend eigenständig sein. Wie in Kapitel 2.3 eingeführt wurde, sind Entitäten Vorkommen in räumlichen Sachverhalten. Einem Vorkommen ist u.a. eine Lokation und eine Identität zugeordnet. In der raumsprachlichen Literatur gibt es Arbeiten,

---

35Dies hat z.T. die Folge, daß dabei entwickelte linguistisch gefärbte Konzepte zur Strukturierung räumlichen Wissens verwendet werden. Dadurch wird eine Nähe zwischen räumlichem Wissen und Sprache suggeriert, die nicht notwendig in einer solch eindeutigen Weise vorhanden ist.
die einerseits die Lokation einer Entität (z.B. [Ta83]) und andererseits die Identität, d.h. die physikalische Ausprägung (z.B. [OB83]), als primär ansehen. Entitäten werden in der deutschen Sprache durch Substantive beschrieben, die im kommunikativen Zusammenhang verschiedene Rollen übernehmen können.

Es ist nahezu unmöglich, Entitäten eines räumlichen Sachverhalts zu beschreiben, ohne gleichzeitig räumliche Relationen zu verwenden. Vielfach werden räumliche Relationen durch räumliche Präpositionen kommuniziert (vgl. [Ta83]). Unter der Annahme, daß räumliche Präpositionen einen direkten Zugriff auf die Struktureigenschaften von räumlichen Relationen erlauben, ist dieser Zusammenhang häufig diskutiert und untersucht worden (z.B. [MIL76], [Ta83], [Her86], [HT95]). In den meisten Arbeiten werden einfach strukturierte räumliche Präpositionen wie „links von“, „über“ und „nahe bei“ betrachtet, von denen angenommen wird, daß sie mittels einer Eins-zu-Eins-Beziehung in räumlichen Relationen verankert sind (z.B. [MIL76], [OB83], [Her86], [Gap94]). Unter Annahme einer dekompositionellen Semantik wird postuliert, daß komplexere Präpositionen, wie beispielsweise „rechts auf“, sich in die Einzelbedeutungen von „rechts“ und „auf“ zerlegen lassen.36

Es gibt zwei Gründe, auf ein Objekt bzw. seine Lokation sprachlich Bezug zu nehmen (vgl. [HG94]):

1. Der Sprecher intendiert, das Objekt hervorzuheben, um die Aufmerksamkeit des Hörers darauf zu lenken.
2. Der Sprecher intendiert, die Lokation hervorzuheben, an der sich das Objekt befindet.

Dies erfolgt auf der Basis von räumlichen Relationen, die durch geeignete sprachliche Ausdrücke, vor allem durch räumliche Präpositionen, kommuniziert werden (vgl. [LJ93], [Kle94], [RS88], [GW95]).

Es lassen sich drei Klassen von räumlichen Präpositionen unterscheiden (vgl. [Her86], [RS88], [Her96]):

1. Projektive Präpositionen
2. Topologische Präpositionen
3. Pfadbezogene Präpositionen


36Ungeklärt ist dabei die Frage, zwischen welchen Relationenklassen Kombinationen möglich sind und zwischen welchen nicht. Im letzteren Fall ist es interessant zu wissen, warum eine bestimmte Kombination nicht möglich ist.
2.5. INTERAKTION VON SPRACHE UND RÄUMLICHES WISSEN

eine bestimmte räumliche Relation erfüllt ist oder nicht. Herskovits bringt das Beispiel, daß
der Satz „Das Buch ist auf dem Tisch“ auch dann noch hinreichend gut erfüllt ist, wenn eine
Tischdecke dazwischen ist. Um dieses formalisieren zu können, wird der sogenannte tolerance
shift eingeführt, der die obige Relation als Eingabe verwendet und nur approximative Gültig-
keit besitzt. Um auf einer geometrischen Ebene solche Abweichungen von der Idealkonfiguration
repräsentieren zu können, wird in verschiedenen Arbeiten das gebietskonstituierende Konzept
einer Anwendbarkeitsfunktion (AF) verwendet (z.B. [ABHR87], [Her96], [Gap94], [OMT94]). Dabei
wird eine räumliche Konfiguration zweier Objekte bezüglich einer räumlichen Relation auf
einen numerischen Anwendbarkeitsgrad (AG) abbildet. Aus der Perspektive der Sprachgenerier-
ung bestimmt der Anwendbarkeitsgrad, wie gut eine dieser räumlichen Relation R zugeordnete
sprachliche Beschreibung die Beziehung zwischen einem Lokalisationsobjekt und einem Refer-
enzobjekt expliziert. Da sich Lagebeziehungen in dynamischen Umgebungen ändern, ist der
Anwendbarkeitsgrad an einen bestimmten Zeitpunkt t gebunden (vgl. [Her86]). Die Anwendbar-
keitsfunktion \(AF_R\) ermöglicht eine lokationsabhängige Bestimmung des Anwendbarkeitswertes
einer räumlichen Relation \(R\) (vgl. [ABHR87], [Gap94], [Her96]). Jede räumliche Relation besitzt
eine charakteristische Anwendbarkeitsfunktion. Über diese Anwendbarkeitsfunktion wird jeder
Lokation des Raumes um das Referenzobjekt ein Anwendbarkeitsgrad zugeordnet. Die Anwend-
barkeitsfunktion wird nicht nur in Abhängigkeit von der Distanz zum Referenzobjekt bestimmt.
Distanzen zwischen den Objekten und die Winkelabweichung des zu lokalisierenden Objektes
von der typischen Lokation bezüglich einer bestimmten räumlichen Relation nehmen entschei-
denden Einfluß auf die Anwendbarkeit einer räumlichen Relation auf die Lage zweier Objekte
(vgl. [Her86, S. 41]).

Eine Anwendbarkeitsfunktion \(AF_R(ego, loc(x), loc(refo), t)\) ordnet jeder Lokation \(x\) im Raum
um das Referenzobjekt \(refo\) zum Zeitpunkt \(t\) in kausaler Referenz zu der Lokation des Be-
trachters \(ego\) einen Anwendbarkeitsgrad \(AG_R\) zu. Die Anwendbarkeitsfunktion berechnet den
Anwendbarkeitsgrad in Abhängigkeit von der Distanz, der Winkelabweichung und den Größen-
verhältnissen zwischen dem Lokalisations- und Referenzobjekt (z.B. [Gap94], [Her96]). In den
bisher entwickelten Modellen ist eine solche Strukturierung einzig auf die geometrischen Aus-
prägungen der Objekte bezogen und berücksichtigt nur geringfügig funktionales oder sonstiges
semantisches Wissen über Objekte.

Das dritte strukturierende Element raumsprachlicher Beschreibungen sind Referenzsysteme. In
der raumsprachlichen Literatur werden Referenzsysteme grundsätzlich danach unterschieden,
ob sie im Sprecher oder in einem Objekt oder einer Lokation außerhalb des Sprechers verankert
sind. Bezogen auf einen kommunikativen Kontext unterscheidet Clark (vgl. [Cla73]), ob ein Spre-
cher eine Umgebung auf sich selbst bezieht (canonical position) oder ob der Rezipient direkt
gegenübersteht (canonical encounter). Hieraus leitet sich eine alternative Unterteilung ab, in
der ein Sprecher oder ein externes Objekt als Zentrum eines Referenzsystems angenommen wird
(z.B. [Her86]). Sprecherbezogene Referenzsysteme werden als egozentriert und objektbezogene
als allozentrische Referenzsysteme bezeichnet. Wie sprachpsychologische Untersuchungen zeigen,
lassen sich allozentrische Referenzsysteme weiter in solche unterteilen, die in Gegenständen und
solche, die in Rezipienten verankert sind (vgl. [Sch95b]). Sprecher zeigten dazu, den kognitiven
Aufwand des Rezipienten für das Verstehen und Verifizieren einer raumsprachlichen Beschrei-

\[37\] In manchen Arbeiten wird der sprachlichen Präposition „zwischen“ eine Sonderrolle zugeordnet, in der dem
Lokalisationsobjekt zwei Referenzsysteme zugeordnet werden (vgl. [ABHR87]). Diese Präposition läßt sich aber auf
raumkognitiver Ebene durch die Komposition von zwei räumlichen Relationen mit jeweils einem Referenzobjekt
modellieren (vgl. [ABHR87]).

\[38\] Empirisch konnten Korrelationen zwischen den Faktoren Distanz, Winkelabweichung und Größe der Objekte
festgestellt werden (siehe [HT95]. Für projektive Relationen gibt es Evidenzen, daß die Distanz keine oder nur
eine untergeordnete Rolle spielt (vgl. [Gap95a]).

\[39\] Der Einfachheit wegen wird t im weiteren Verlauf weggelassen.
bung zu verringern (vgl. [Her89], [HGH91], [Sch95b]). Dabei versetzt sich der Sprecher meist in die Position des Rezipienten, was in der vis-à-vis-Anordnung einen maximalen kognitiven Aufwand bezüglich mentalen Rotationen bei der Generierung intrinsischer Beschreibungen bedeutet (vgl. [HBN87]). Dennoch werden solche rezipientenzentrierte Beschreibungen präferiert (vgl. [Meh65], [Som69]).


Wie räumliches Wissen in sprachlichen Beschreibungen, insbesondere Wegbeschreibungen, verwendet wird, ist Gegenstand der folgenden Abschnitte.

### 2.5.1 Raumdeixis

Raumsprachliche Beschreibungen lassen sich in solche unterscheiden, die sich auf das unmittelbare Umfeld des Sprechers beziehen und solche, deren Referenzen darüber hinausgehen. Referenzen auf unmittelbare Umgebungen konstituieren das, was als Zeigfeld (vgl. [B34]) oder auch Deixis bezeichnet wird (vgl. [Kle79]). Allgemein lassen sich die Deixis des Raums und der Zeit unterscheiden (vgl. [Lev89], [Kle79]). Raumdeiktische Ausdrücke wie „hier“ und „dort“ referenzieren auf solche Entitäten, die sowohl vom Sprecher, als auch vom Rezipienten in der aktuellen Situation identifizierbar sind. Dreierei kann ein Sprecher durch raumdeiktische Ausdrücke vermitteln: identifizieren, informieren und bestätigen (vgl. [Fil82]). Im Hinblick auf die Verankerung von sprachlichen Ausdrücken in visuellen Daten, sind die ersten beiden Aufgaben von deiktischen Beschreibungen zu beachten. Im Kontext der Situationssemantik (vgl. [BP83]) werden solche Sprachausdrücke mit Referenten in der Situation durch Referenzfunktionen verknüpft. Die Aufgabe des Hörers ist es, die Instantierung der Referenzfunktion des Sprechers zu antizipieren (vgl. [Nun79]). Ein referentieller Ausdruck kann bereits vorher eingeführt worden sein, oder es besteht für den Rezipienten die Aufgabe, ein geeignetes Objekt in der Situation zu bestimmen. Um die Unifikation von Referenz und potentiellen Referenten zu erleichtern, ist es die Aufgabe eines kooperativen Sprechers, die Beschreibung möglichst eindeutig zu halten. Dies kann u.a. durch Angabe von visuellen Merkmalen erfolgen, die in der Situation als salient erscheinen. Dabei muß zwischen solchen Merkmalen unterschieden werden, die der Sprecher selbst und solche, von denen der Sprecher annimmt, daß der Hörer sie als salient empfindet. Letztendlich ist es die Aufgabe des Sprechers, eine Lokation oder ein Objekt durch geeignete Beschreibungen derart zu referenzieren, daß der Rezipient sie eindeutig identifizieren kann (vgl. [Lev89]).


---

40Im weiteren Verlauf kurz als deiktische Ausdrücke bezeichnet.

41In der Situationssemantik wird die Gesamtheit aller dieser Referenzen als die Verbindungen des Sprechers (speaker's connections) bezeichnet (vgl. [BP83]). Eine Situation im Sinne der Situationssemantik ist gleichbedeutend mit dem deiktischen Raum, der den Sprecher und den Hörer gemeinsam umgibt.
2.5. INTERAKTION VON SPRACHE UND RÄUMLICHEM WISSENS

2.5.2 Diskursbereich der Wegbeschreibungen

Wegbeschreibungen sind ein im Alltag häufig vorkommender Diskurstyp, der sprachliche, perzephuelle und mentale Imaginationsfähigkeiten mit räumlichem Wissen verbindet. Wegbeschreibungen sind sowohl in ihrer kommunikativen Form, als auch inhaltlich stark schematisiert (vgl. [Kle82], [Wun79]). Dies erlaubt es dem Sprecher, Information effizient auszuwählen und zu kommunizieren, solange er annimmt, daß der Rezipient über ähnliche Schemata verfügt. Solche Kommunikationsschemata reglementieren sowohl das, was durch eine kommunikative Handlung an Information ausgetauscht werden kann, als auch die Art und Weise, wie dies prototypisch erfolgt.\footnote{Nach Herrmann und Grabowski kann das für das Sprechen erforderliche Wissen in Was-Wissen und Wie-Wissen unterteilt werden (vgl. [HG94]). Mit Was-Wissen wird der Diskurstyp charakterisiert. Um welchen Diskurstyp es sich handelt, wird entweder explizit verhandelt, ist durch die Situation vorgegeben oder wird implizit durch die Art und Weise, wie einer der beiden Kommunikationsteilnehmer das Gespräch beginnt, vorgegeben (vgl. [Sch87]). Wie-Wissen gibt vor, wie beztiglich eines Diskurstyps eine Beschreibung strukturiert wird. Soweit das Was- oder Wie-Wissen erlernt ist, werden sie zu Routinen integriert, die Herrmann et al. Was- und Wie-Schemata bezeichnen (vgl. [HG94], für empirische Untersuchungen siehe [HKDD02]). Beispiele für die Verwendung von solch schematisiertem Sprachwissen sind Wohnungsbeschreibungen (vgl. [LL75]).}


Im weiteren Verlauf wird auf den Diskurstyp Wegbeschreibung eingegangen und dabei insbesondere auf die Linearisierung. Als wesentliche Aufgabe von Wegbeschreibungen werden anschließend die referentiellen Ausdrücke genauer diskutiert. Abschließend werden relevante formale Modelle vorgestellt.

2.5.2.1 Wegbeschreibungen als Diskurstyp

Wegbeschreibungen sind aus linguistischer Sicht vielfach untersucht worden (vgl. [Kle79], [Wun82], [Hab87], [MMP^88], [HCR90]), was auf die eindeutige kommunikative Aufgabenstellung zurückzuführen ist.\footnote{In manchen Arbeiten werden anstelle von „Wegbeschreibung“ auch andere Begriffe wie „Routeninformation“ oder „Wegauskunft“ verwendet, die aber im wesentlichen synonym verwendet werden können.} Das Ziel einer Wegbeschreibung ist es, einem Rezipienten einen Weg von einem Startort zu einem Zielort sprachlich und durch Gesten zu beschreiben. Typischerweise hat der Auskunftsgewgende alleine das relevante räumliche Wissen bezüglich der Problemlösung.

Von einer kommunikationsorientierten Perspektive aus gehören Wegbeschreibungen zur Klasse der Wegauskünfte (vgl. [Kle79]). Danach besteht eine Wegauskunft aus einer Einleitung, einer Zentralsequenz (Wegbeschreibung) und einem Abschluß. Wunderlich und Reinhart haben die Zentralsequenz weiter in eine Phase der Wegbeschreibung und einer Absicherungsphase unterteilt (vgl. [WR82]). Diese Phasen sind in ihrer Ausführung nicht sequentiell, sondern interagierend miteinander.

Der Diskurstyp, der durch die Phase der Wegbeschreibung festgelegt ist, läßt sich in die beiden Klassen der vollständigen und inkrementellen Wegbeschreibungen untergliedern (vgl. [Maa93]). Durch eine vollständige Wegbeschreibung beschreibt der Auskunftgeber den Weg ohne direkte Referenzierung auf die beschriebene Umgebung. Der Weg wird vom Sprecher mental verfolgt. Das Ziel einer vollständigen Wegbeschreibung ist es somit, beim Rezipienten eine solche räumlich-zeitliche Vorstellung zu erzeugen, die es ihm erlaubt, den wesentlichen Informationsgehalt der
Beschreibung zu memorieren und später eindeutig mit der Real-Umgebung abzugleichen.44 Drei Elemente stellt Klein als wesentliche Bestandteile einer vollständigen Wegbeschreibung heraus. Das erste Element sind Fixpunkte (Landmarken oder Kreuzungen). Relativ zu Fixpunkten können Richtungen und Aktionen referenziert werden. Die Aufgabe von Wegbeschreibungen liegt darin, beim Hörer einen deiktischen Raum zu erzeugen, der mit dem des Sprechers hinreichend übereinstimmt. Da vollständige Wegbeschreibungen inhärent dynamisch sind, d.h. daß die Position des Sprechers bzw. des Rezipienten sich permanent ändert, modifiziert sich auch der deiktische Raum in dem Sinne, daß er beim Rezipient erweitert wird. Um zu erkennen, was der Sprecher als Fixpunkt bezeichnet, stützt sich der Rezipient auf vorhandenes Wissen, visuell perzipierbare Objekte sowie deiktische und relative Beschreibungen zu bereits eingeführten Fixpunkten.

Die Grundlage vollständiger Wegbeschreibungen ist das intern45 oder extern46 memorierte räumliche Wissen des Sprechers über eine Umgebung. Soll ein Weg beschrieben werden, so aktiviert der Sprecher einen Ausschnitt dieses Wissens aus der Kognitiven Karte, welcher den Start- und Zielpunkt umfaßt. Dieser Ausschnitt wird als Primärplan bezeichnet. Der Primärplan wird weder vollständig im voraus noch schrittweise entwickelt (vgl. [Kle79]). Klein postuliert, daß der Primärplan zuviel Information enthält, als daß er in seiner Vollständigkeit dem Hörer beschrieben werden kann. Demzufolge wird der Primärplan überarbeitet, d.h. es wird nur relevante Information ausgewählt. Das Ergebnis ist ein sogenannter Sekundärplan, der als direkte Eingabe für die Sprachgenerierung dient (vgl. [Kle79]).


---


45Wissen, welches im Gedächtnis gespeichert wird.

46Wissen, welches durch die Verwendung externer Medien, wie beispielsweise Karten, inferiert werden kann.
2.5.2.2 Linearisierung durch reale und imaginierte Wanderungen

Nach Levelt gibt es zwei grundlegende Determinanten, nach denen holistisches, räumliches Wissen über eine Domäne sequenzialisert werden kann. Zum einen legt das Prinzip der natürlichen Ordnung eine inhaltbezogene Determinante fest, nach der zu beschreibende Information so angeordnet wird, wie sie die natürliche Ordnung des Inhalts angibt (vgl. [Lev89, S. 138]). Die prozeßbezogene Determinante kommt dann zum Tragen, wenn keine inhaltbezogene Determinante anwendbar ist. Sie wird durch verschiedene Prinzipien beeinflußt, wie beispielsweise das Prinzip der Verbindungheit (vgl. [Lev89, S. 139ff.]).\(^{47}\)

Abbildung 2.11: Linearisierung einer Wegbeschreibung durch den sich ändernden Perzeptionsbereich

Im Fall des Diskurstyps der Wegbeschreibungen ist die natürliche Ordnung durch die intendierten Trajektorien des Weges und die assoziierte Struktur der imaginierten oder direkt perzipierten Umgebung determiniert. Bei inkrementellen Wegbeschreibungen perzipiert der Auskunftgeber zu einem Zeitpunkt jeweils nur eine Situation. In dieser Situation wird eine begrenzte Menge an Landmarken ausgewählt, die entsprechend ihrer Bedeutung für die Beschreibung in die sprachliche Inhaltsstruktur integriert werden. Dazu werden solche Landmarken ausgewählt, die, ähnlich dem Ankerpunktprinzip (vgl. [CGGT87]), zu dem aktuellen Verzweigungspunkt in Beziehung stehen (siehe Abbildung 2.11). Die Linearisierung erfolgt somit entsprechend der sequentiellen Abfolge von Verzweigungspunkten entlang eines Pfades. Bezüglich Verzweigungspunkt VP\(_2\) perzipiert der Agent die Landmarken L\(_1\) und L\(_2\), d.h., daß beide im Perzeptionsbereich liegen. Im

Perzeptionsbereich fokussiert der Sprecher auf einen Bereich, der um den nächsten Verzweigungs-
punkt liegt (räumlicher Aufmerksamkeitsbereich). Der Sprecher präferiert solche Landmarken,
die im räumlichen Aufmerksamkeitsbereich liegen, da sie eine räumliche Nähe Verzweigungs-
punkt besitzen. Demzufolge wird nur L_{2} ausgewählt, da L_{1} von VP_{2} zu weit entfernt ist. Dies illustriert, daß die Beschreibung des Wegabschnitts von VP_{1} nach VP_{2} alleine durch eine lokale Beschreibung einer Aktion am Verzweigungspunkt angezeigt wird. Somit reduziert sich das Li-
nearisierungsproblem im Kontext von Wegbeschreibungen auf ein solches für die Beschreibung einer Aktion an einem Verzweigungspunkt. Wie die empirischen Daten zeigen (siehe Kapitel 3), werden bei inkrementellen Wegbeschreibungen Aktionen in vielen Fällen durch Referenzen auf Landmarken und Zeitmarkern disambiguier. Beschreibungen von Aktionen, Landmarken und Zeitmarkern können im Deutschen in flexibler Weise permutiert werden, wodurch die Sequentia-
isierung der Beschreibung auf grammaticalik Regeln reduziert ist. Die Hierarchisierung des
perzierten Raums determiniert über das Prinzip der Verbundenheit die globale Sequenzi-
sierung der Beschreibung (siehe Abbildung 2.11; Aktionen an VP_{1} werden vor solchen an VP_{i+1}
beschrieben). Auf der lokalen Ebene eines Verzweigungspunktes werden sprachliche Ausdrücke
entsprechend grammaticalischer Abhängigkeiten und persönlichen Präferenzen sequenziert.

In ähnlicher Weise wie bei inkrementellen Wegbeschreibungen erfolgt die Sequenziierung bei
vollständigen Wegbeschreibungen (vgl. [Kle79], [WR82]). Anstelle einer realen Bewegung durch
eine Umgebung wird das Konzept einer imaginiären Wanderung (vgl. [LL75], [Kle79]) oder eines
generischen Wanderers\(^{18}\) (vgl. [HG94])) verwendet. Entsprechend der Metapher der imaginiären
Wanderung kann sich der Auskunftgebende in seiner Kognitiven Karte in ähnlicher Weise wie in
einer Realumgebung „umschauen“. Bisher sind aber erst Ansätze einer Theorie erkennbar,
die beschreiben, welche Funktionen und Repräsentationen es erlauben, räumliche Information
aus einer Kognitiven Karte zu extrahieren (siehe Kapitel 2.3.7). Sprachorientierte Arbeiten zu
vollständigen Wegbeschreibungen basieren auf der Hypothese, daß die Strukturierung und Se-
quentialisierung des memorierten räumlichen Wissens auf ähnliche Weise wie beim Aufbau räum-
lchen Wissens via direkter Perzeption erfolgt. Aus diesem Grunde verwenden solche Ansätze
auch eine Strukturierung entsprechend der sequentiellen Abfolge der Verzweigungspunkte. Wie
Verzweigungspunkte und Landmarken referenziert werden, wird im weiteren diskutiert.

2.5.2.3 Referentielle Ausdrücke in Wegbeschreibungen

Die Qualität einer Wegbeschreibung ist um so höher, je geringer die Ambiguitäten bei gleichzei-
tiger Minimierung des kognitiven Verstehensaufwandes beim Hörer sind. Da der Sprecher keinen
direkten Zugriff auf den kognitiven Aufwand beim Rezipienten besitzt, ist ein heuristisches Maß
durch die Länge der Wegbeschreibung gegeben. Dadurch entsteht ein „trade-off“ zwischen der
Disambigierungsfähigkeit und der Länge einer Wegbeschreibung (vgl. [Gri75]).

Eine sprachliche Referenzbildung auf ein Objekt erfolgt selten durch Aufzählung aller Merkmale
(vgl. [Lev89, S. 129]). Vielmehr wählt der Sprecher aus den gememorierten oder perzipierten Merk-
malen solche aus, die er dem Rezipienten erläuern, das Objekt eindeutig zu identifizieren.
Wie psycholinguistische Studien zeigen, sind referentielle Ausdrücke nahezu immer ausreichend
oder nicht-ambivalent. Jedoch enthalten sie meist redundante Elemente (für einen Überblick
siehe [DP82]). Weiterhin werden referentielle Ausdrücke bevorzugt, die leichter zu verstehen
sind als solche, die einfacher zu beschreiben sind (vgl. [RD92]). Reiter und Dale beschreiben ein
Modell, in dem das präferierte Merkmal eines Objektes (z. B. Größe, Form und Farbe) bevorzugt
wird, falls es die Zahl der referenzierbaren Objekte reduziert.

---

\(^{18}\) Die Verwendung des Begriffs „Wanderung“ ruht daher, daß diese Untersuchungen meist mit Fußgängern
durchgeführt worden sind. Jedoch lassen sich entsprechende Untersuchungen auch unter Verwendung anderer
Bewegungsarten durchführen.
2.6. MODELLER ALLGEMEINER RAUMKOGNITIVER LEISTUNGEN


Generell setzt die Verbalisierung eines referentiellen Ausdrucks voraus, daß der Sprecher und Rezipient hinreichend ähnliche Grundannahmen haben. Nicht immer ist dies gegeben, so daß es bei nicht erfolgreich interpretierter referentieller Ausdrucks auf Seiten des Sprechers zur Aktivierung von Reparaturmechanismen kommt (vgl. [Edm93, S. 17]). In Bezug auf referentielle Ausdrücke werden Kommunikations situatiorien unterschieden, in denen der Sprecher und der Rezipient gemeinsames Wissen über die referenzierten Objekte besitzen (vgl. [App85a, AK87, HHe91, Se69]) oder in denen die referenzierten Objekte im Aufmerksamkeitsbereich beider liegen (vgl. [RD92, GS86]).

Bei der Auswahl eines geeigneten Referenzobjektes zur Beschreibung einer Lokation spielen vor allem Salienzfehler eine Rolle (vgl. [PC90]). Die Salienz eines Objektes ist von der Betrachtungsperspektive, der Bekanntheit und von den Zielen des Sprechers bzw. Rezipienten abhängig (vgl. [Lyn60]). Über die Perspektive sind perzeptuelle Eigenschaften, wie Identifizierbarkeit, Sichtbarkeit und Auffälligkeit von Einfluß (vgl. [Dev76]). Das Konzept der Salienz ist aber besonders durch die perzeptuelle Auffälligkeit beeinflußt, welche wiederum nur relativ zu einer Situation bestimmt werden kann.⁴⁹

2.6 Modelle allgemeiner raumkognitiver Leistungen

Eine Vielzahl formaler Modelle sind zur Verarbeitung räumlichen Wissens entwickelt worden. Dabei werden zumeist entweder einzelne raumkognitive Leistungen isoliert oder mehrere integriert betrachtet.

1. Isolierte raumbezogene Leistungen

(a) Objektrepräsentationen
   (z.B. [Bin71], [Neu89], [Koe90], bei der Verarbeitung visueller Daten [Bro81, Mar82, Rob63, Low87, HU90])

(b) Räumliche Relationen
   (z.B. [RS88], [Gap94], [Han80], [CJ85], [ABHR87] [AK93], [Ege91], [RCC92], [Sch9a, OMT94], [Her91], [Zim93], [R94])

(c) Erkennen und Verarbeiten von Ereignissen
   (z.B. [NN83a], [Her92], [Nag88], [PB95])

(d) Sprachliche Raumausdrücke
   (z.B. [ABHR87], [RS92], [BSZ87], [Pri93])

⁴⁹ Appelt nennt den ersten Fall “non-shared concept activation with identification intention” (vgl. [App85c]) und den zweiten “shared concept activation with identification intention” (vgl. [App85c]). Im Zusammenhang mit inkrementellen Wegbeschreibungen werden referentielle Ausdrücke des zweiten Typs verwendet. Vollständige Wegbeschreibungen hingegen verwenden nahezu ausschließlich solche des ersten Typs.

⁵⁰ Dieser Punkt wird häufig hervorgehoben, jedoch nahezu immer aufgrund seiner vermuteten Komplexität nur unzureichend, d.h. statisch behandelt (siehe z.B. [Edm93]).
2. Integrierte raumbezogene Leistungen
   (vgl. [Kui77], [MS79], [Fun81], [EL82], [SVW85], [KL88], [GKS89], [LZ89], [Dav86], [MD84],
   [Yan88], [HW94])

Die erste Sichtweise betrachtet einzelne Bestandteile eines raumkognitiven Systems und entwickelt hierzu formale Modelle. Die zweite Sichtweise stellt die Frage, wie verschiedene raumkognitive Leistungen integriert werden können, um ein bestimmtes Problem in einer komplexen Umgebung zu lösen. In diesem Sinne sind beide Richtungen komplementär. Im weiteren werden Ansätze zur Lösung isolierter Leistungen nicht weiter diskutiert (hierzu vgl. [Her86]). In Bezug auf das hier vorgeschlagene Modell eines Agenten ist die zweite Kategorie von primärem Interesse.

2.6.1 Das TOUR Modell

In den ersten Arbeiten zur Modellierung raumkognitiver Leistungen durch Computermodelle ist meist eine Zusammenstellung zwischen isolierten und integrierten Leistungen eingenommen worden (vgl. [Kui77], [MS79], [EL82]). Kuipers TOUR-Modell ist in der Lage, räumliches Wissen durch einen Assimilationsprozeß zu lernen und für Wegfindungszwecke zu verwenden (vgl. [Kui77], [Kui78]). Dieses Modell basiert auf entwicklungspychologischen Theorien (vgl. [Pls60], [SW75]) und den räumlichen Konzepten von Lynch (vgl. [Lyn60]). In den ersten Arbeiten unterschied Kuipers nur zwischen prozeduralem und topologischem Wissen (vgl. [Kui78]). Im Laufe der Zeit erweiterte er dieses Modell zur "spatial semantic hierarchy", die verschiedene Wissenssebenen unterscheidet (vgl. [Kui94]). In der sensomotorischen Ebene ist das gesamte räumliche Wissen des Agenten verankert. Die Information dieser Ebene steht in enger Beziehung zur Steuerungsebene, auf der Steuerungsgesetze zur Trajektorienverfolgung auf lokaler Merkmalsebene definiert sind. Die kontinuierliche Information der sensomotorischen und der Steuerungsebene wird auf diskreten Strukturen abgebildet, die zum einen Sichten und zum anderen Aktionen enthalten. Sichten repräsentieren Integrationen von sensorischer und aktionsorientierter Information zur Trajektorienverfolgung. Auf der Kausalebene werden zwei Sichten (z.B. V und V') und eine Aktion (A) zu Assoziationen integriert (\(A < V, A, V' >\)). Sichten werden meist an Verzweigungspunkten generiert, wohingegen Aktionen solche sind, die ein Roboter benötigt, um von einer Lokation, die einer Sicht zugeordnet ist, zur nächsten zu gelangen. Dementsprechend korreliert dies mit dem Übergang von Landmarken zu Routenwissen (siehe Kapitel 2.3.5). Auf der topologischen Ebene werden Sicht-Aktion-Assoziationen zu Plätzen, Pfaden und Regionen integriert, zwischen denen topologische Relationen, wie Verbindung und Inklusion etabliert werden. Dies korrespondiert zum Übergang von Routen- und Landmarkenwissen zu Konfigurationswissen (siehe 2.3.6). Erst nach dem Aufbau topologischen Wissens kann metrisches Wissen, wie Distanzen, Richtungen, Formen u.a. abgeleitet werden.

In dem Modell, welches durch die räumlich-semantische Hierarchie gegeben ist, werden im Vergleich zu anderen Modellen die meisten Arten räumlichen Wissens integriert. Dennoch bleiben einige Fragen offen. Beispielsweise ist nicht geklärt, ob die Unterteilung in schnappschußartige Sichten und verbindende Aktionen auch in komplexen Situationen eine geeignete Konzeptualisierung darstellt. Ebenso ist es möglich, metrische Information direkt von sensorischer Information abzuleiten, ohne eine topologische Repräsentation aufzubauen. Die Sequenzialisierung bei der Entwicklung von Landmarken- und Routenwissen, so wie sie von Piaget vorgeschlagen

\[\text{Das in [KL88] beschriebene Modell wird um die Steuerungsebene erweitert.}\]

\[\text{Obwohl dies von Kuipers nicht explizit beschrieben wird, ist anzunehmen, daß Sichten nur beim Stillstand des Roboters etabliert werden können. Dies führt in dynamischen Situationen, in denen der Roboter permanent in Bewegung bleiben muß, zu unakzeptablen Leistungen.}\]
2.6. MODELL ALLGEMEINER RAUMKOGNITIVER LEISTUNGEN

Es werden genauer betrachtet, finden in der psychologischen Literatur Gegenbeispiele (vgl. [Bla91]). Eine mögliche Erweiterung besteht somit in der Auffassung dieser sequenziellen Architektur.


2.6.2 Das Modell ELMER


2.6.3 Das Modell von Elliott und Lesk


2.6.4 Das Modell TRAVELLER


---

53 Aus einer ingenieursorientierten Perspektive ist die räumlich-semantische Hierarchie nicht zu beanstanden, solange sie ihren Zweck erfüllt. Küppers postulierte jedoch auch eine kognitionswissenschaftliche Relevanz seiner Arbeit, weswegen sie sich auch an dieser zu messen hat.

Die Behauptung von Leiser und Zilbershatz, daß der TRAVELLER im Hinblick auf eine kognitive adäquate Repräsentation des Wissens einer kognitiven Karte bessere Eigenschaften als das TOUR-Modell besitzt, wird nicht erfüllt. Im Gegensatz zum TOUR-Modell basiert das Lernen des Routenwissens auf einer Rasterrepräsentation. Dieses enthält inhärente metrische Beziehungen, von denen sich andere räumliche Relationen direkt ableiten lassen. Das topologische Konzept der räumlichen Nähe ist eine triviale Folgerung daraus. Dadurch kehrt das TRAVELLER-Modell die Entwicklung räumlichen Wissens in die Ableitung topologischem von metrischem Wissen um, was empirischen Befunden widerspricht (siehe Kapitel 2.3).

2.6.5 Das Modell NAVIGATOR


Abbildung 2.12: Schematische Darstellung der funktionalen Einheiten des NAVIGATOR-Modells nach Gopal (vgl. [GKS89])

Das NAVIGATOR-Modell umfaßt eine zweidimensionale, rasterförmige Umgebung und das eigentliche *kognitive Modul* (NV). NV bewegt sich durch von außen vorgegebene Anweisungen von

---

55Es sind, ähnlich wie auf der topologischen Ebene des TOUR-Modells, nur Translationen und Rotationen um 90° Winkel möglich.

Weiterhin verfügt der NV über Funktionen des Vergessen. Dazu wird der Salienzwert einer Entität im Arbeitsgedächtnis proportional zur Zeit schrittweise vermindert.57 Bei wiederholter Perzeption einer Entität wird der Salienzwert erhöht.58 Hat NV in einer Umgebung ausreichendes räumliches Wissen aufgebaut, ist er in der Lage, sich auf Anfrage zu einer Lokation zu bewegen, die der beschriebenen Situation am nächsten kommt.59 Durch mustergesteuerten Vergleich sucht NV im Arbeits- und Langzeitgedächtnis eine geeignete Szene. Der Aufbau raumbezoge-

---

56 Die interne Repräsentation räumlichen Wissens ist mit dem Format des externen, räumlichen Wissens identisch.
57 Die entsprechende Funktion des Salienzwertes in Abhängigkeit von der Zeit lautet: $S(t) = S_0 \cdot e^{-\beta(t-t_0)}$, wobei $t_0$ die Zeitpunkt der Perzeption einer Entität und $\beta$ ein globaler Parameter ist.
58 Die Erhöhung des Salienzwertes bei n-maliger Perzeption lautet: $\delta S_n = S_0 \cdot \alpha^n$, mit der Konstanten $\alpha < 1$. Hieraus folgt, daß die Gesamtsalienz $S_n$ eines Objektes sich wie folgt berechnet: $S_n = \sum_{n=0}^{\infty} S_0 \cdot \alpha^n = \frac{S_0}{1 - \alpha}$.
59 Eine Beispielanfrage ist: „Gehe zu einem roten Haus mit einem blauen Auto und einem hölzernen Zaun neben dem Auto.“ (vgl. [Gop88]).

Die Unterscheidung in eine objektive Umgebung und ein kognitives Modul erscheint für die Untersuchung raumbezogener kognitiver Fähigkeiten als geeigneter Ansatz, um die Dynamik der involvierten Prozesse und Repräsentationen besser untersuchen zu können. Letztendlich stellt das Modell einen interessanten Versuch der Modellierung des Prozesses des räumlichen Wissenserwerbs dar.

2.6.6 Das Basic Agent-Modell


Zusammenfassend läßt sich sagen, daß der Beitrag, der durch das Modell erbracht worden ist, in der Integration komplexer Leistungen besteht. Der Schwerpunkt liegt in Homer auf den Planungs- und Schlußfolgerungsmechanismen und der Anbindung an NL-Mechanismen.

2.6.7 Das Modell SPAM

Das von McDermott entwickelte SPAM²³-Programm ist als autonome Wissensbasisverwaltung für räumliches Wissen konzipiert worden (vgl. [MD84]). SPAM besteht aus drei Modulen. Die Integrationseinheit integriert räumliches Wissen in einer Kognitiven Karte, durch welche per Anfrage über zwei Anfrageeinheiten auf einzelne oder eine Liste von Objekten zugegriffen werden kann. Zur Lokation eines Objektes ist SPAM ein Referenzsystem zugeordnet. Relationen

²³Spatial module


2.6.8 Das Modell MERCATOR

Das MERCATOR-Modell baut durch Erkundung einer Umgebung eine Kognitive Karte auf (vgl. [Dav86]). In diesem Modell werden drei Wissensstrukturen unterschieden:

1. Wissen über Dimensionen, Orientierungen und Winkel innerhalb eines Objektes sowie zwischen Objekten

2. Wissen über Objektformen

3. Wissen über die Größe eines Objektes


2.6.9 Das Modell von Yeap


### 2.6.10 Zusammenfassung raumverarbeitender Modelle


Im folgenden wird die Unterklasse raumverarbeitender Modelle diskutiert, die Funktionalitäten zur Generierung von Wegbeschreibungen formalisieren.
## Tabelle 2.3: Systemvergleich: Indizierte Modellierung entspricht Funktionalitäten; o indizierte Modellierung entspricht Partial-Funktionalitäten

<table>
<thead>
<tr>
<th>Modelle</th>
<th>TOUR</th>
<th>ELMER</th>
<th>SOCCER</th>
<th>Elliot</th>
<th>NX</th>
<th>TRAVELLER</th>
<th>NAVIGATOR</th>
<th>Basic Agent</th>
<th>MERCATOR</th>
<th>SPAM</th>
<th>Yesp</th>
<th>MOSES*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agentenkonzept</td>
<td>o</td>
<td>o</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D-Umgebungen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-Umgebungen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Realumgebungen</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geometrische Repräsentation</td>
<td>o</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visuelle Datenerfassung</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berechnung visueller Salienzen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natürlichsprachliche Eingabe</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natürlichsprachliche Ausgabe</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adaptiv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“anytime” Verhalten</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple Objektrepräsentationen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topologische Relationen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordnungsrelationen</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metrische Relationen</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landmarkenwissen</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Routenwissen</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konfigurationswissen</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egozentrische Referenzsysteme</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allozentrische Referenzsysteme</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Globale Referenzsysteme</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kognitive Karten</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Name des implementierten Agenten des in Kapitel 4 beschriebenen funktionalen Modells
2.7 Modelle zur Generierung von Wegbeschreibungen

Zur sprachlichen Analyse von Wegbeschreibungen gibt es empirische Studien, theoretische (vgl. [Hab87], [HK91], [HCR90], [Con95]) und implementierte Berechnungsmodelle (vgl. [RS87], [MMP+88], [Müll88], [Car91], [Dav89]). Diese setzen sich ausschließlich mit vollständigen Wegbeschreibungen auseinander. Theoretische Modelle sind konzeptionelle Beiträge, die hauptsächlich Funktionen identifizieren, von denen angenommen wird, daß sie bei der Generierung von Wegbeschreibungen eine Rolle spielen. Implementierte Modelle legen den Schwerpunkt auf die Operationalisierbarkeit und Implementation solcher Funktionen.\textsuperscript{62}

2.7.1 Das formale Modell von Habel


\textsuperscript{62}Es wird nur eine Auswahl von Modellen diskutiert, da sich andere Modelle auf diese zurückführen lassen (wie z.B. [Mar95], [Gry92], [May92], [BB03]).

Operationalisierung dieser einzelnen Schritte an.

Abbildung 2.14: Interne Struktur eines Routenauskunftssystems nach Habel (vgl. [Hab87])


„... da, wo das Cafe Bley ist, wo man draußen sitzen kann ...“ (vgl. [Hab87, S. 37])


2.7.2 Das formale Modell von Höök und Karlgren

Vor einem mehr ingeniurwissenschaftlichen Hintergrund untersuchen Höök und Karlgren die Kommunikationsstruktur von Wegbeschreibungen (vgl. [HK91]). Dazu unterscheiden sie zwei Gruppen von Rezipienten: Touristen und Stadtteilinwohner. Durch empirische Untersuchungen


### 2.7.3 Das formale Modell von Hoeppner et al.

2.7.4 Das formale Modell von Coucellis

Coucellis (vgl. [Cou95]) stellt den Rahmen eines Wegbeschreibungssystems vor, in dem die Konzepte eines mentalen Modells (vgl. [IL83], ähnlich zu Lakoffs “idealized cognitive model” (vgl. [Lak87]) und Fauconniers “mental spaces” (vgl. [Fau85])) mit dem Konzept der Kognitiven Karte (vgl. Kapitel 2.3.7) verbunden werden. Sie nimmt an, daß zur Generierung von Wegbeschreibungen das in mentalen Modellen repräsentierte spezifische und generelle Wissen über Umgebungen mit intentionalen Zuständen und Verhaltensformen integriert ist. Kognitive Karten, so postuliert Coucellis, sind ebenso wie Wegbeschreibungen in mentalen Modellen der Wegbeschreibungssituation verankert. Dies widerspricht der allgemeinen Annahme, daß Kognitive Karten die Grundlage von raumsprachlichen Beschreibungen sind. Mentale Modelle selbst sind wiederum in primitiven, kognitiven Strukturen, insbesondere “image-schemas” (vgl. [Ta83], [Lak87]) und “basic-level” Kategorien (vgl. [Ros73]), verankert und repräsentieren darüber die gesamte Wegbeschreibungssituation.

![diagram](image)

Abbildung 2.15: Die Stufen zur Generierung von Wegbeschreibungen (nach [Cou95])


eigenständige, zwischen Sprache und Raum vermittelnde Repräsentationsformate, sondern neben raumbezogenen Diskursen als ein mögliches Explikationsformat aufzufassen. Es bleibt offen, in welcher Form räumliches Wissen in solcher Art mentalen Modellen repräsentiert wird und ob damit Phänomene der Kognitiven Karten (z.B. Verzerrungen) erklärt werden können.

2.7.5 Das Modell von Rau und Schweitzer


2.7. MODELLE ZUR GENERIERUNG VON WEGBESCHREIBUNGEN

Abbildung 2.16: Systemarchitektur nach Rau und Schweitzer (nach [RS87])


2.7.6 Das Berechnungsmodell KLEIST


Regeln
Propositionen
NS Komponente
Ausgabe
Umwandlung der Zustände in Aktionen
Strukturierung der VS-Liste
Liste Zust. B. (VS)
Erstellung einer Liste von Zustandsbeschreib.
Liste von WAs
Wegfindung
Analyse der Eingabe
Verhalten
Eingabe
Marken
Pfade
Zustands Sprit
Regeln Scripts
Regeln Propos.


### 2.7.7 Das Berechnungsmodell CITYGUIDE

CITYGUIDE ist eines der ersten vollständig implementierten Modelle, welches eine natürlichsprachliche Beschreibung eines Weges ermöglicht (vgl. [Mü88], [Wah91]). Wie vergleichbare Ansätze (vgl. [HCR90], [Car91]), basiert dieses Modell auf einer netzartigen Repräsentation der räumlichen Umgebung.

Kante mit allen Knoten verbunden ist, die Endpunkte von in den Platz einmündende Straßen repräsentieren. Jeder dieser Endknoten ist untereinander verbunden.\footnote{Es bleibt unklar, weshalb eine solche komplexe Repräsentation eingeführt wurde. Die Zusammenfassung aller Einmündungsknoten in einen Platz erhält offenbar ebenfalls die in CITYGUIDE betrachteten Erfordernisse.}


Die klare Aufteilung in einen Primär- und Sekundärplan, so wie sie von Klein vorgeschlagen wurde, wird in CITYGUIDE nur rudimentär durch Zusammenfassung von Pfadabschnitten zu Routen über die Indizierung durch den Straßennamen berücksichtigt. ROUTENABSCHNITTE werden durch einfache Konkatenation miteinander verbunden, was dazu führt, daß längere Beschreibungen sehr künstlich und monoton klingen. Unverständlich werden die Beschreibungen sogar dann, wenn Rezipienten, Straßen- oder Landmarkennamen genannt werden. Ebenso nicht berücksichtigt werden lokallekische und temporale Ausdrücke. Da CITYGUIDE sich ausschließlich auf Daten bezieht, die dem Wegnetz entstammen, wird die Beschreibung, metaphorisch gesehen, aus einer Vogelperspektive beschrieben. Routenwissen wird nicht verwendet, was auch dadurch zum Ausdruck kommt, daß Merkmale von Landmarken nur einen konstanten Auffälligkeitswert besitzen.

### 2.7.8 Das Berechnungsmodell von Carstensen


Abbildung 2.18: Das Modell der Generierung von Wegbeschreibungen (nach [Car91])

Das Modell ist eine Anwendung der 2-Stufen-Semantik (vgl. [BL87]) auf den Bereich der Wegbeschreibungen. Hierzu wird eine konzeptuelle Ebene eingeführt. Konzeptuelle Strukturen besitzen einerseits eine vermittelnde Rolle zwischen der perzeptuellen und andererseits der motorischen Ebene und der sprachlichen Ebene. Jedoch bleibt Carstensen besonders für die Verbindung...

2.7.9 Das Back Seat Driver Modell


67 Der TravelPilot von Blaupunkt ist nur eins von vielen Beispielen.
<table>
<thead>
<tr>
<th>WB-Systeme</th>
<th>Habel</th>
<th>Hook</th>
<th>Hoeppner</th>
<th>Condelis</th>
<th>Rau</th>
<th>KLEIST</th>
<th>CITY GUIDE</th>
<th>Carstensen</th>
<th>Davis</th>
<th>MOSES*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vollständige Wegbeschreibungen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Inkrementelle Wegbeschreibung</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Primär/Sekundärplan</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Geometrische Umgebung</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>“dual coding” Repräsentation</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vollständiges räumliches Wissen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Limitiertes räumliches Wissen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egozentrisches Referenzsystem</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Allozentrisches Referenzsystem</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imaginäre Wanderung*</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Berechnung räumlicher Relationen</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Gradierte räumliche Relationen</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verschiedene Wegetypen</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verschiedene Rezipiententypen</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landmarken</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Feste visuelle Salienz</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berechnete visuelle Salienz</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plätze</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Räumliche Nähe</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schemabasierte Sprachgenerierung</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empirische Studien</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementiert</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Name des implementierten Agenten des in Kapitel 4 beschriebenen funktionellen Modells

*Imaginäre Wanderung
2.7.10 Zusammenfassung der Wegbeschreibungsmodelle


Vor diesem Hintergrund bieten inkrementelle Wegbeschreibungen einen Ansatzpunkt, zu untersuchen, wie unmittelbar perzipierte Information verwendet wird, um räumliches Wissen aufzubauen und sprachlich zu verwenden. Zusätzlich sind aber auch zeitliche Aspekte zu betrachten, die durch die Bewegung des Sprechers und der Umgebung selbst induziert werden.

2.8 Zusammenfassung

In diesem Kapitel sind Elemente indiziert worden, die in einer vollständigen Theorie der Raumkognition enthalten sein müssen. Bisher ist nur schemenhaft zu erkennen, wie eine solche Theorie aussehen kann. In der empirischen Literatur herrscht weitgehend Einigkeit darüber, daß es amodale, eigenständige Repräsentationen räumlichen Wissens gibt, die aus unterschiedlichen Perzeptionsfunktionen abgeleitet werden. In der sprachwissenschaftlichen Literatur herrscht diese Einigkeit nicht in einer solch eindeutigen Weise. In einflussreichen Ansätzen (z.B. [Jac87], [Her86]) wird implizit oder explizit die Auffassung vertreten, daß räumliches Wissen Teil des grammatischen-lexikalischen Wissens ist. Im Bereich der Künstlichen Intelligenz gibt es Arbeiten, die die eine oder andere Perspektive einnehmen. Letztendlich kann diese Frage bisher nicht mit endgültiger Sicherheit beantwortet werden, was nicht zuletzt auf der Komplexität räumlichen Wissens beruht.


\textsuperscript{55} Ausgenommen ist hiervon das TOUR Modell.
Kapitel 3

Empirische Untersuchungen zu inkrementellen Wegbeschreibungen


In Bezug auf inkrementelle Wegbeschreibungen sind empirische Untersuchungen bisher nicht durchgeführt worden. Aus diesem Grund ist als Startpunkt im Rahmen dieser Arbeit durch eine empirische Untersuchung einen Sprachkorpus inkrementeller Wegbeschreibungen erstellt und analysiert worden. Da inkrementelle Wegbeschreibungen alltägliche kommunikative Handlungen sind, ist die Untersuchung in einer Realumgebung durchgeführt worden, was eine hohe ökologische Validität gewährleistet. Bedingt durch die geringe Kontrollierbarkeit von Situationseinflüssen in der ersten Untersuchung, ist in einer zweiten unter Laborbedingungen eine einzelne Verkehrs situation verwendet worden, um Auswirkungen kontrollierter Parametervariationen auf das Sprachverhalten zu überprüfen. Im weiteren wird nach einigen Vorbemerkungen die Felduntersuchung und anschließend die Laboruntersuchung beschrieben.

3.1 Vorbemerkungen

Die Kommunikationshandlung, die durch inkrementelle Wegbeschreibungen gegeben ist, wirft eine Reihe empirischer Fragen auf. Aus linguistischer Sicht sind vollständige Wegbeschreibungen häufig auf ihre grammatische Struktur untersucht worden (vgl. Kapitel 2.5.2). Solche Beschreibungen beruhen auf Situationen, die vom Sprecher auf Grund seiner Erfahrung erinnert werden. Die Aufgabe der Probanden besteht in solchen Untersuchungen darin, von memoriertem, raumbezogenem Wissen Beschreibungen abzuleiten. Bei der Generierung vollständiger Wegbeschreibungen sind in erheblichem Maße Phänomene zu berücksichtigen, die auf Inference-
schritten des Sprechers über den Rezipienten beruhen (z.B. [HK91], [Kle79]). Da der Rezipient eine mentale Vorstellung alleine auf der Basis der Beschreibungen ableitet, sind diese vom Sprecher derart zu strukturieren, daß sie in ausreichendem Maße memoriert werden können und später verifizierbar sind (vgl. [HCR90]). Es vermischt sich in vollständigen Wegbeschreibungen somit das allgemeine Problem der Generierung von Raumbeschreibungen mit solchen, die sich auf Memorierungsleistungen beziehen. In vergleichbaren Arbeiten ist untersucht worden, welches raumbezogene Wissen Probanden bei einer Autofahrt durch eine fremde Umgebung aufbauen (z.B. [ALM64], [CS69]). Dabei zeigt sich eine große Übereinstimmung bei den memorierten Objekten (vgl. [CS69]), was die kommunikative Verwendbarkeit von solchen Objekten erhöht. Im Kontext der inkrementellen Wegbeschreibungen verifiziert der Rezipient die aus einer Beschreibung abgeleitete Bedeutung direkt in der perzipierbaren Umgebung, so daß der Aufwand der Gedächtnisleistungen erheblich verringert wird.

Maßgeblich für die Generierung inkrementeller Wegbeschreibungen ist die Eigenbewegung des Sprechers und des Rezipienten. Damit verknüpft sind Abschätzungen über Geschwindigkeiten und die Verwendung anderer wahrnehmbarer Eigenschaften einer Umgebung. Vor diesem Hintergrund lassen sich für inkrementelle Wegbeschreibungen (IWB) folgende Hypothesen aufstellen:

1. Inhaltsstrukturen von IWB sind aus einer geringen Menge von sprachlichen *Inhaltskategorien* aufgebaut.
2. Inhaltliche und sprachliche Strukturen von IWB sind an Zeitbeschränkungen der Situation angepaßt.
3. Die Referenzierbarkeit eines Objektes setzt eine ausreichend hohe visuelle Salienz voraus.
4. Objekte werden nur dann referenziert, wenn sie sich an geeigneten räumlichen Lokationen befinden.


---

¹In einer diagonal hierzu angelegten Untersuchung haben Streeter und Kollegen ([SVW85]) die Verwendung verschiedener Informationsmodi als Fahrerangabeninformationen verglichen. Von einem Kassettenrekorder inkrementell gegebene Anweisungen sind durch den Versuchsleiter vorgegeben worden. Sie fanden heraus, daß verbale Beschreibungen effektiver vom Rezipienten verwendet werden können als eine Straßenkarte oder einer gleichzeitigen Verwendung einer verbalen Beschreibung und einer Straßendatei.

²Dynamische Landmarken, wie beispielsweise der Halesche Komet, werden selten verwendet.

³Die Erweiterung durch 'K' indiziert, daß es sich um eine Inhaltskategorie handelt.
3.1. VORBEREHRKUNGEN


Zusammenfassend lassen sich die vier Inhaltsstrukturen wie folgt charakterisieren:

1. WOHIN-K: indiziert die Richtung der auszuführenden Aktion
2. WO-K: referenziert über die Beschreibung eines Objektes oder eines Pfadelementes ein Gebiet
3. WANN-K: gibt eine qualitative4 Beschreibung des Zeitpunktes zur Ausführung einer Aktion
4. WAS-K: referenziert die Art der Bewegung

Im Bereich der Forschung zur visuellen Wahrnehmung sind Aufmerksamkeits- und Fokussierungsprozesse in zahlreichen Studien untersucht worden. Der Einfluß solcher Aufmerksamkeitsprozesse auf Prozesse der sprachlichen Ebene ist hingegen bisher weniger untersucht worden (z.B. [GS86, AK87]). Eine Fragestellung, die aus einer integrierten Betrachtungsweise von visueller Wahrnehmung und Sprache entsteht, ist, in welcher Weise visuell fokussierte Entitäten in sprachlichen Beschreibungen verwendet werden, aber auch, ob und wie sprachliche Prozesse Aufmerksamkeitsprozesse beeinflussen (vgl. [BJKZ85]). Solche Fragestellungen führten zu empirischen Untersuchungen (z.B. [Coh84], [CM95]) und formalen Modellierungen (z.B. [App85b], [RD92], [HH92], vgl. Kapitel 2.5). Eine der Fragen in diesem Kontext betrifft die Auswahl der Kriterien für visuell perzipierte Objekte. In der Aufmerksamkeitsforschung werden momentan

---


Im folgenden werden zwei empirische Untersuchungen vorgestellt. In der ersten empirischen Untersuchung wurden die Probanden aufgefordert, in einer realen Umgebung einem Rezipienten inkrementelle Wegbeschreibungen zu geben. In der zweiten Untersuchung mußten die Probanden Beschreibungen bezüglich einer computeranimierten Kreuzungssituation verbalisieren.

### 3.2 Explorative Untersuchung in realen Umgebungen


#### 3.2.1 Arbeitsmodell

Als Arbeitsmodell wird angenommen, daß die Generierung von inkrementellen Wegbeschreibungen eine konventionalisierte, kommunikative Handlung ist. Dies bedeutet, daß sowohl der Sprecher, als auch der Rezipient schematisches Wissen über den prototypischen Ablauf einer inkrementellen Wegbeschreibung besitzen. Dies erlaubt es dem Sprecher, mit minimalem sprachlichen Aufwand die Information zu übermitteln, von der er/sie ausgeht, daß der Rezipient sie benötigt, um den Pfad eindeutig zu finden (vgl. [Kle79], [Wun78], [HG94]). In vollständigen Wegbeschreibungen besteht der größte Aufwand für den Rezipienten darin, die Beschreibungen in ausreichender Weise zu memorieren, um sie später zu verifizieren (vgl. [HCR90]). Da die In-
formation inkrementeller Wegbeschreibungen direkt in der aktuellen Situation verwendet wird, ist zu vermuten, daß inkrementelle Wegbeschreibungen kürzer und prägnanter sind. Weiterhin wird postuliert, daß die Detailliertheit einer Beschreibung von der Zeitdauer bis zum Erreichen eines Verzweigungspunktes abhängt.

### 3.2.2 Durchführung


Eine Beispieltranskription, welche zur Beschreibung des Pfadabschnitts 8 gegeben wurde, ist wie folgt:

"... und hier vorne jetzt an der Kreuzung links.\ (10 sec) Abschnitt 8"

Dieses Transkript gibt neben der Beschreibung selbst an, daß nach der vollständigen Verbalisierung des Satzes 10 Sekunden vergingen, bis der Rezipient den Verzweigungspunkt am Ende des Pfadabschnitts 9 erreicht hat.

### 3.2.3 Ergebnis

Insgesamt wurden 142 einzelne Beschreibungen aufgenommen. Eine Analyse der Transkripte ergab, daß inkrementelle Wegbeschreibungen aus vier Grundkategorien und einer Restkategorie aufgebaut sind, die unterschiedlich häufig verwendet wurden (vgl. Abbildung 3.2). In die Restkategorie wurden Aussagen, wie z.B. „Vorsicht Stufe!“ aufgenommen. Insgesamt betrug der Anteil der Beschreibungsteile, die in die Restkategorie aufgenommen wurden, 26% der Fälle. Ausdrücke der Restkategorie sind beispielsweise „... so ...“ „... und ...“ sowie „... okay, also ...“. Diese Ausdrücke markieren in den meisten Fällen den Beginn einer neuen Beschreibung bzw. eine Gedankenpause. Da solche Ausdrücke direkt keine räumliche Information enthalten, werden sie im weiteren Verlauf nicht eingehender betrachtet.


\footnote{Diese relativen Häufigkeiten sind von den Anteilen der Restkategorie bereinigt.}
schwacher Zusammenhang zu erkennen, der darauf beruht, daß kombinierte Inhaltskategorien aus einzelnen Inhaltskategorien zusammengesetzt sind. Beispielsweise läßt sich aus den relativen Häufigkeiten der einzelnen Inhaltskategorien WO-K und WOHIN-K nicht erklären (vgl. Abbildung 3.2), warum die relative Häufigkeit der Kombination von WO-K und WOHIN-K größer als die der WOHIN-K Kategorie ist (vgl. 3.3).


Eine Analyse der Zeitpunkte, an denen inkrementelle Wegbeschreibungen gegeben worden sind, ergibt eine Häufung von Beschreibungen kurz vor dem Erreichen eines Verzweigungs punktes (Zeitraum T_5 in Abbildung 3.4). Je weiter der Proband noch vom Verzweigungspunkt entfernt ist, um so gleichmäßiger sind die Beschreibungen zeitlich verteilt. Je näher der Sprecher
dem nächsten Verzweigungspunkt (VP) ist, um so kürzer werden die Beschreibungen. Inhaltlich wird in einem größeren zeitlichen bzw. räumlichen Abstand zum VP das Gebiet um den VP durch WO-K Instanzen genauer beschrieben. Kurz vor Erreichen von VP wird die auszuführende Aktion deiktisch durch eine zeitliche oder räumliche Referenz (vor allem durch „hier“ und „jetzt“) markiert. Der Sprecher geht davon aus, daß das Gebiet um den VP hinreichend genau vom Rezipienten erkannt worden ist.

Sprachlich zeigt sich dieser Wechsel dadurch, daß in einem großen raum-zeitlichen Abstand vollständige, wohlgeformte Sätze gebildet werden, in denen räumliche und zeitliche Referenzen sowie Aktionen genau spezifiziert werden. Bei Annäherung an den Verzweigungspunkt werden nur noch elliptische Beschreibungen gegeben, die zumeist auf deiktischen Referenzen beruhen. Die Analyse der Transkripte bezüglich der Häufigkeiten von inkrementellen Wegbeschreibungen
erbrachte fünf Zeitrahmen: \( T_1, \ldots, T_5 \). Bedingt durch sprachliche und inhaltsliche Unterschiede der Beschreibungen zeigt sich, daß die einzelnen Zeitrahmen nicht äquidistant sind, sondern um so kürzer werden, je näher sie dem nächsten VP sind. Die nachfolgende Aufführung zeigt Beispiele für Beschreibungen, die den einzelnen Zeitrahmen zugeordnet werden.

1. \( T_1 \): „Und vor dem Musikbau da vorne auf der linken Seite gehts rechts die Treppe hoch. Also zwischen Post und Musikgebäude“ (85 sec. bis zum Erreichen des VP)
2. \( T_2 \): „Und am roten Baum da vorne wieder rechts und dann die Treppe hoch“ (35 sec. bis zum VP)
3. \( T_3 \): „... und da an der Kreuzung nach rechts, da so den Berg rauf.“ (15 sec. bis zum VP)
4. \( T_4 \): „... so und hier an der Kreuzung gehts jetzt links.“ (10 sec. bis zum VP)
5. \( T_5 \): „Hier rechts.“ (2 sec. bis zum VP)

Beschreibungen, die sich \( T_1 \) zuordnen lassen, rekurrieren meist auf Erfahrungen der Probanden über die Umgebung. Solche Effekte werden im weiteren nicht weiter behandelt. Beschreibungen innerhalb der Zeitrahmen \( T_2 \) bis \( T_5 \) hingegen sind nahezu ausschließlich durch Referenzen auf visuell perzipierbare Entitäten und Merkmale erfolgt. Nur in fünf Ausdrücken wurde insgesamt auf den Namen eines Gebäudes referenziert. Diese Einteilung in Zeitrahmen ist eine erste Approximation. In weiteren Untersuchungen muß analysiert werden, inwieweit es sich um diskrete oder kontinuierliche Prozesse handelt. Eine genaue Analyse zeigt, daß die Probanden, mit Ausnahme von einem, bei dieser Einteilung während der Zeitrahmen \( T_3 \) bis \( T_5 \) eine nahezu identische relative Häufigkeit besitzen, was durch die geringe Standardabweichung bezüglich der einzelnen Zeitrahmen von 16\% indiziert wird.

<table>
<thead>
<tr>
<th>Zeiträume</th>
<th>Dauer bis zum VP (sek)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_5 )</td>
<td>1-3</td>
</tr>
<tr>
<td>( T_4 )</td>
<td>3-10</td>
</tr>
<tr>
<td>( T_3 )</td>
<td>10-25</td>
</tr>
<tr>
<td>( T_2 )</td>
<td>25-40</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>&gt; 40</td>
</tr>
</tbody>
</table>

Tabelle 3.1: Zuordnung der Zeiträume zu Zeitintervallen

Die meisten Beschreibungen werden während des Zeitrahmens \( T_4 \) gegeben. Diese Tendenz ist, bis auf eine Versuchsperson, typisch. Während \( T_2 \) und \( T_3 \) differieren die Häufigkeiten zwischen den Personen stark, was daran liegt, daß die Beschreibung des Gebiets um den VP in verschiedenen Weise erfolgte. Beschreibungen zum Zeitrahmen \( T_2 \) sind sehr ausführlich, wobei während \( T_3 \) bereits eine Fokussierung erfolgt.

Eine Zuordnung des Inhalts einer Beschreibung zu den Zeitrahmen ist durch Rückgriff auf die Inhaltskategorien möglich. Dabei zeigt sich, daß die Häufigkeit von Instanzen einzelner Inhaltskategorien während den Zeiträumen \( T_2 \) und \( T_3 \) in der gleichen Größenordnung liegen (vgl. Abbildung 3.5). Während des Zeitrahmens \( T_4 \) kommt es zu einem parallelen Anstieg in der Verwendung von Instanzen der \( WANN-K \) und \( WAS-K \) Kategorien. Im Zeitrahmen \( T_5 \) geht die Verwendung von Instanzen der beiden Kategorien wieder auf ein Minimum zurück. In ähnlicher Weise verläuft die Verwendung von Instanzen der \( WOHN-K \) und der \( WO-K \) Kategorien parallel. Nach einem Minimum im Zeitrahmen \( T_3 \) steigt die Häufigkeit an und erreicht in \( T_5 \) ein Maximum. Dies indiziert, daß Instanzen der \( WANN-K \) und \( WAS-K \) Kategorien dazu verwendet werden, Aktionen
frühzeitig anzukündigen. In dem Augenblick, in dem die Aktion ausgeführt werden soll, wird die räumliche Lokation (WOHIN-K) zusammen mit einer Richtungsangabe (WOHIN-K) beschrieben.

Abbildung 3.6: Relative Häufigkeiten der Kombination von Inhaltskategorien bezüglich einzelner Zeitrahmen

Die Verwendungshäufigkeit einzelner Kategorien hat Auswirkungen auf die Inhaltsstruktur der gesamten Beschreibung. In Abbildung 3.6 sind die relativen Häufigkeiten von Kategorienkombinationen dargestellt. Prinzipiell lassen sich zwei Tendenzen identifizieren: Je kürzer die Zeit bis zum Erreichen eines Verzweigungspunktes ist, um so stärker steigt die relative Häufigkeit, mit der die WOHIN-K + WO-K Kombinationen verwendet werden. Ein inverses Verhalten zeigt sich für die Verwendung der anderen Kombinationen, die auf Erweiterungen von WAS-K + WOHIN-K Kombinationen beruhen und ebenso für die Kombination WOHIN-K + WANN-K + WO-K.
### 3.2. Explorative Untersuchung in realen Umgebungen

<table>
<thead>
<tr>
<th></th>
<th>WAS-K</th>
<th>WOHN-K</th>
<th>WANN-K</th>
<th>WO-K</th>
<th>relative Häufigkeit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_2$</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,23</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,15</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,46</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,15</td>
</tr>
<tr>
<td>$T_3$</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,25</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,25</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,38</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,13</td>
</tr>
<tr>
<td>$T_4$</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,18</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,12</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,24</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,30</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,12</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,03</td>
</tr>
<tr>
<td>$T_5$</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,58</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,05</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,07</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,05</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,15</td>
</tr>
<tr>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>0,07</td>
</tr>
</tbody>
</table>

Tabelle 3.2: Zeitrahmenbezogene relative Häufigkeiten von W*-Kategoriekombinationen

#### 3.2.4 Diskussion


Die Aufteilung in Zeitrahmen korrespondiert mit der Unterscheidung in verschiedene subjektive Räume (z.B. [Itt73], [Man83], [GG87], [Mon93]). Obschon der Proband die aktuelle Situation in $T_2$ und $T_3$ überblickt, befindet er sich noch innerhalb des rein visuellen Raums. Entsprechend der Einteilung in Zeitrahmen erstreckt sich der visuelle Raum von $T_2$ über $T_3$ bis hin zu $T_4$. Mit Übergang zum Zeitrahmen $T_5$ erfolgt ein Übergang vom visuellen Raum zum interaktiven Raum (Handlungsraum). Dieser Übergang manifestiert sich in der häufigen Verwendung deiktischer Referenzen. Vollständige Wegbeschreibungen auf der Basis räumlicher Langzeitrepräsentationen des Sprechers basieren auf einem ähnlichen Übergang (vgl. [Kle79], [Wun78]). Da der Übergang nicht auf perzipierter Information basiert, sondern imaginiert erfolgt, korrespondiert dies mit dem Konzept der imaginären Wanderung (vgl. [LL75]). Der Sprecher deutet dabei dem Rezipienten durch Verwendung von deiktischen Beschreibungen an, daß das Gebiet erreicht worden ist, an dem eine Aktion erfolgt. Dies bewirkt, daß der Rezipient seine Aufmerksamkeit auf das beschriebene Gebiet richtet, um die verbalsprachlich gegebenen Referenzen mit geeigneten Referenten zu identifizieren und die Aktion auszuführen. Dies erfolgt typischerweise durch redundante Beschreibungen eines einzelnen Sachverhalts. Eine solche Verwendung redundanten Beschreibungen zeigt sich auch in dem hier aufgebauten Korpus der inkrementellen Wegbeschreibungen. Dabei zeigt eine Analyse, daß redundante Beschreibungen charakteristisch verteilt gegeben werden. Wird eine Beschreibung zum Zeitraum $T_1$ gegeben, so wird eine redundante Beschreibung zum Zeitraum $T_{1+2}$ gegeben.


Weiterhin ist zu untersuchen, welchen Einfluß die Variation der Bewegungsgeschwindigkeit des Probanden auf die Verwendung von Inhaltskategorien und deren Kombinationen besitzt. Insbesondere im Hinblick auf die Analyse von Ressourcenbeschränkungen können solche Untersuchungen Hinweise liefern.

### 3.3 Untersuchung in einer computeranimierten 3D-Umgebung

In Beschreibungen visuell wahrgenommener, räumlicher Umgebungen referenziert ein Sprecher nicht auf alle Details. Dies entspricht vor allem kognitiven und kommunikativen Ökonomiekriterien (vgl. [Gri75]). Eine Teilaufgabe, die ein Sprecher während der inhaltlichen Festlegung einer inkrementellen Wegbeschreibung zu lösen hat, besteht darin, solche Elemente aus einer Umgebung auszuwählen, die eindeutig identifizierbar sind. Eine solche referenzierende Rolle übernehmen Landmarken (vgl. Kapitel 2.3.4). In makroskopischen, stadtähnlichen Umgebungen werden häufig Gebäude als Landmarken verwendet (z.B. [Lyn60], [ALM64]). Daß Landmarken eine wichtige Rolle beim Erwerb und der Beschreibung raumbezogenen Wissens spielen, ist vielfach diskutiert worden (vgl. Kapitel 2.3.4). Offen bleibt in diesen Arbeiten, auf Grund welcher Kriterien ein Objekt als Landmarke verwendet wird. Gleiches gilt für situationssemantische Mo-
3.3. UNTERSUCHUNG IN EINER COMPUTERANIEMITZEN 3D-UMGEBUNG

delle, in denen zwischen sprachlichen Ausdrücken und referenzierten Objekten Verbindungen ("connection") postuliert werden (vgl. [BP83]).


3.3.1 Arbeitsmodell

Das Arbeitsmodell, welches der empirischen Untersuchung zugrunde liegt, postuliert, daß sowohl durch visuelle Merkmale, als auch durch räumliche Lokationen beeinflußt wird, ob ein Objekt in einer inkrementellen Wegbeschreibung als Landmarke verwendet wird. Damit verbunden ist die Frage, wodurch die unterstützende Referenzierung visueller Merkmale beeinflußt wird. Bezogen auf die einzelnen visuellen Merkmale wird vermutet, daß das Merkmal Farbe den größten Einfluß besitzt (vgl. [Man86]). Weiterhin wird postuliert, daß die visuelle Salienz eines Objektes sich aus einer gewichteten Linear-Kombination visueller Merkmale zusammensetzt, und daß die Verwendung eines Objektes als Landmarke von pfadbasierten Intentionen abhängt.

3.3.2 Durchführung


Die Untersuchung bestand aus vier Teilen (vgl. Tabelle 3.3). In den ersten beiden Teilen wurde

---

6In dieser Untersuchung ist der Fall nicht betrachtet worden, daß Landmarken beschrieben werden, um die Lokation einer Kreuzung von anderen Kreuzungen hervorzuheben. In diesem Falle stellt sich die Frage, wie ein primärdeiktischer Raum unabhängig von der intendierten Bewegungsrichtung etabliert wird. Da die Richtung der Bewegung am nächsten Verzweigungspunkt nur eine untergeordnete Rolle spielt, ist zu vermuten, daß zur Identifizierung der Kreuzung auch Landmarken auf der „falschen“ Seite verwendet werden.

7Eine eingehende Diskussion der allgemeinen Wirkung von Merkmalen auf Klassifikations- und Sprachprozesse findet sich in [Kli71].
Abbildung 3.7: Kreuzungssituation mit den Pfadabschnitten $S_i$, der Kreuzung $C$ und den Gebäuden $L_i$

<table>
<thead>
<tr>
<th>Freier Fall</th>
<th>Festgelegter Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farbe</td>
<td>Gruppe 1</td>
</tr>
<tr>
<td>Höhe</td>
<td>Gruppe 2</td>
</tr>
</tbody>
</table>

Tabelle 3.3: Zuordnung von Probandengruppen zu Untersuchungsvariationen

die Farbe genau einer Landmarke variiert und in den anderen beiden deren Höhe. Um überprüfen zu können, inwieweit sprachliche Phänomene bei der Referenzierung von Landmarken wesentlich sind, wurde die Variation der Farbe und der Höhe in zwei Teile geteilt. Im ersten konnten die Probanden die Situation frei beschreiben (freier Beschreibungsfall), wohingegen sie im zweiten Teil aufgefordert wurden, eine bestimmte, über eine Nummer indizierte Landmarke in die Beschreibung zu integrieren (festgelegter Beschreibungsfall).

Die Probanden wurden in zwei Gruppen à sechs Personen geteilt. Die erste Gruppe beschrieb die Situation mit variierenden Farben im freien Modus und mit variierender Höhe im festgelegten Modus. Im Gegensatz dazu beschrieb die zweite Gruppe Situationen mit variierenden Höhen im freien und variierende Farben im festgelegten Modus.

Für beide Variationen erhielten die Probanden die Instruktion, sich vorzustellen, als Beifahrer in einem Auto zu sitzen. Ihre Aufgabe bestand darin, so natürlich wie möglich einem neben ihnen sitzenden Rezipienten verbal zu beschreiben, wohin dieser zu fahren hat. Der Proband wurde darüber informiert, daß kurz vor Beginn der animierten Situation am unteren Rand ein Pfeil eingeblendet wird, der die Richtung der Aktion am nächsten Verzweigungspunkt angibt. Ein Pfeil nach oben indizierte eine Geradeausfahrt, ein Pfeil nach links einen Linksabbiegevor-

---

8Da die Probanden sich nicht bewegen sollten, ist diese Anweisung natürlicher, als wenn sie sich vorstellen, ein Fußgänger zu sein.
3.3. UNTERSUCHUNG IN EINER COMPUTERANIMIERTEN 3D-UMGEBUNG

<table>
<thead>
<tr>
<th>Farbe</th>
<th>H/L/S 1</th>
<th>H/L/S 2</th>
<th>H/L/S 3</th>
<th>H/L/S 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>grau</td>
<td>000/075/000</td>
<td>000/065/000</td>
<td>000/055/000</td>
<td>000/045/000</td>
</tr>
<tr>
<td>rot</td>
<td>029/052/058</td>
<td>029/042/065</td>
<td>029/032/085</td>
<td>029/025/100</td>
</tr>
<tr>
<td>blau</td>
<td>180/054/014</td>
<td>180/044/015</td>
<td>180/034/019</td>
<td>180/024/017</td>
</tr>
</tbody>
</table>

Tabelle 3.4: Tabelle der varierten Farbwerte (H), Helligkeiten (L) und Sättigungsgrade (S)

gang, respektive ein Pfahl nach rechts einen Rechtsabbiegevorgang. Durch Tastendruck konnte

der Proband die jeweils nächste Situation starten. Im festgelegten Fall mußten die Probanden
zusätzlich die Landmarke in die Beschreibung integrieren, deren Nummer am unteren Rand des
Bildschirms erschien. Die Referenz der Nummer auf die Landmarke wurde durch eine Karte
hergestellt, die den Probanden vorlag (vgl. Abbildung 3.7).

Die graphischen Animationen der Kreuzungssituation sind in LISP mit CLIM-Funktionen imple
mentiert worden (vgl. Kapitel 4.3.2). Das Animationsprogramm auf der Basis der Programmpakete GEOMETRIX und GEODISPLAY (vgl. Kapitel 4.3) lief auf einer Workstation von Hewlett Packard (HP 9000/720) und zeigte die Animationen auf einem 19"-Farbmonitor (Hewlett Packard HP98754A) an. Der Proband interagiert mit dem Animationssystem über den
Druck der Leerzeichen-Taste.

3.3.2.1 Methode: Farbvariation

Im ersten Fall der Variation des Farbmerkmal einer Objektes wurden die Farbmerkmale der an
deren Objekte konstant hell-grau (HLS: 000/065/000) gehalten. Alle Objekte hatten identische
geometrische Ausmaße. Variiert wurde der Farbwert, die Helligkeit und der Sättigungsgrad, die
Lokation und die Richtung der zu beschreibenden Aktion. Als Farbgrundwerte wurden exem
phatisch grau, rot und blau mit vier verschiedenen Helligkeits- und Sättigungsgraden verwendet
(vgl. Tabelle 3.4).

3.3.2.2 Ergebnis: Farbvariation

Wurde ein Objekt im freien Beschreibungsfall sprachlich referenziert, erfolgte die Nennung des
Farbmerkmals über Farbwerte, Helligkeits- und Sättigungsvariationen, Lokationen des Objektes
und Aktionen hinweg im Durchschnitt mit einer relativen Häufigkeit von 14% (Standardabweie
chung von 5% über alle Farbvariationen). Im festgelegten Beschreibungsfall kam es zu einer hohen
Standardabweichung von 17% über alle Farbvariationen. Im Durchschnitt lag die Verwendung
von Farbmerkmal zur Beschreibung einer Landmarke bei 32% (vgl. Abbildung 3.8).

Durchschnittlich wurde im freien Beschreibungsfall mit gleicher Häufigkeit eine Landmarke ohne
(10% der Fälle) und mit Indikation ihres Farbmerkmals (14% der Fälle) referenziert. Im fest
glegeten Fall: hingegen wurde in 19% der Fälle häufiger eine Landmarke ohne (51%) als mit
Nennung des Farbmerkmals (32%) beschrieben. Diese Unterschiede blieben über Farbwerte und
Helligkeits- und Sättigungsgradvariationen hinweg konstant (vgl. Abbildung 3.8).

Ein deutscher Unterschied in der Referenzierung des Farbmerkmals ergibt sich beim Vergleich
von Beschreibungen von Abbiegeaktionen und Geradeausfahrten. Bei Geradeausfahrten wurde
im freien Beschreibungsfall keine Landmarke durch ihr Farbmerkmal genauer spezifiziert. Im fest-
glegeten Fall hingegen wurde in Beschreibungen mit Objektreferenzen bei einer Geradeausfahrt

---

9hue, lightness, saturation
10Ein spezieller Farbabgleich ist nicht durchgeführt worden, da es in erster Linie um die Untersuchung qualita
tiver, nicht quantitativer, Unterscheidungen ging.
mit einer Häufigkeit von 46% auf das Farbmerkmal referenziert. Dies ändert sich grundsätzlich bei intendierten Abbiegeaktionen. Im freien Fall wurden in 73% der Beschreibungen mit Landmarken (Standardabweichung 9%), bzw. in 25% aller Beschreibungen (Standardabweichung 3%) das Farbmerkmal genannt. Im festgelegten Fall verringerte sich dies auf 48% (Standardabweichung 6%). Dies zeigt, daß Probanden vor allem bei Abbiegeaktionen Farbmerkmale mit ausreichender visueller Salienz referenzierten. Der festgelegte Fall zeigt, daß zwar ein Farbmerkmal bei einer Geradeausfahrt auffällig ist, es jedoch nie verwendet wird.

Die räumliche Lokation der Landmarke ist in den Beschreibungen von herausragender Bedeutung. Waren die Probanden frei in ihrer Beschreibung, so verwendeten sie für die Beschreibung einer Abbiegeaktion ausschließlich (100%) solche Landmarken, deren Lokation auf der Abbiege-seite („richtige“ Seite, vgl. Abbildung 3.9) lag. Auch wenn eine Landmarke auf der zur Aktion abgewandten Seite („falsche“ Seite) farblich besonders auffällig war, wurde sie in der Beschreibung nicht berücksichtigt.

### 3.3.2.2.1 Diskussion: Farbvariationen


Abbildung 3.9: Qualitative Unterscheidung in Landmarken, die bei einem Linksabbiegevorgang auf der „richtigen“ (A und B) und solche, die auf der „falschen“ Seite (X und Y) liegen.

### 3.3.2.3 Methode: Höhenvariation


### 3.3.2.4 Ergebnis: Höhenvariation

Wurde eine Landmarke im freien Beschreibungsfall sprachlich referenziert, erfolgte die Nennung des Merkmals Höhe über Höhenwert-, räumliche Lokations- und Aktionsvariationen hinweg im Durchschnitt mit 17% bei einer Standardabweichung von 17% über alle Höhenvariationen. Diese hohe Streuung ist dadurch bedingt, daß die Referenzierung des Höhenmerkmals mit ansteigendem Höhenwert bis auf 36% erheblich ansteigt. Im festgelegten Fall wurde ein Höhenmerkmal in 51% der Fälle bei einer Standardabweichung von 4% über alle Höhenvariationen verwendet (vgl. Abbildung 3.10). Dies zeigt, daß das Merkmal Höhe im festgelegten Fall über alle Höhenvariationen hinweg gleich verwendet wird. Im freien Beschreibungsfall wurde eine Landmarke in nur 2% aller Fälle referenziert und dabei in 17% dieser Fälle zusätzlich durch ihr Höhenmerkmal genauer spezifiziert. Im festgelegten Fall verschob sich dieses Verhältnis auf 5% zu 51%.

Ein signifikanter Unterschied zur Referenzierung des Farbmerkmals ergibt sich beim Vergleich der Beschreibungen von Abbiegeaktionen und Geradeausfahrten. Bei Geradeausfahrten wurde im freien Beschreibungsfall keine Landmarke durch ihr Höhenmerkmal genauer spezifiziert. Im festgelegten Fall wurde relativ zu allen Beschreibungen über Höhenwerte hinweg bei einer Geradeausfahrt konstant mit einer Häufigkeit von 89% auf das Höhenmerkmal referenziert.
Abbildung 3.10: Relative Häufigkeiten für das Merkmal *Höhe* im freien und im festgelegten Beschreibungsfall bei verschiedenen Höhen $I_i$

Dies ändert sich deutlich bei intendierten Abbiegeaktionen. Im freien Fall wurden in 53% aller Beschreibungen (Standardabweichung 46%) bzw. in 93% der Beschreibungen, in denen Landmarken (Standardabweichung 11%) referenziert wurden, das Höhenmerkmal genannt. Die große Streuung kommt wiederum dadurch zustande, daß über die Höhenwerte hinweg die Referenzierung des Höhenmerkmals von 8% bis auf 100% zunimmt. Im festgelegten Fall erhöht sich die Referenzierungshäufigkeit auf 97% (Standardabweichung von 5%).

Der identische Effekt wie bei der Variation von Farbmerkmalen gilt auch hier für die räumliche Lokation der varierten Landmarke. Waren die Probanden frei in ihrer Beschreibung, so wandten sie für die Beschreibung eines Abbiegevorgangs ausschließlich (100%) Landmarken, die auf der „richtigen“ Seite lagen. Auch wenn eine Landmarke auf der „falschen“ Seite besonders hoch war, wurde sie in einer Beschreibung nicht berücksichtigt.

3.3.2.4.1 Diskussion: Höhenvariation Aus den Ergebnissen folgt, daß bei einem Abbiegevorgang mit Referenz einer Landmarke in 93% aller Beschreibungen sie durch ihr Höhenmerkmal genauer spezifiziert wurde. Dieser Effekt ist dabei weitgehend davon unabhängig, wie hoch die Landmarke ist. Vielmehr mußte die Höhe nur einen Schwellenwert überschreiten. Im festgelegten Fall steigt die Verwendungshäufigkeit auf 97% an. Ebenso wie in der vorhergehenden Untersuchung des Farbmerkmals ist für die Verwendbarkeit einer Landmarke in einer Aktionsbeschreibung deren räumliche Lokation von großem Einfluß. Der Sprecher integriert ausschließlich Landmarken, die auf der „richtigen“ Seite bezüglich einem Abbiegevorgang liegen.

Bei einer Geradeausfahrt spielt im freien Beschreibungsfall das Höhenmerkmal keine Rolle. Wurde der Proband jedoch aufgefordert, diese Landmarke zu integrieren, so wurde sie in 89% aller Fälle und in 97% bei einem Abbiegevorgang zusätzlich durch ihr Höhenmerkmal spezifiziert.

3.3.3 Gesamtdiskussion


Einen starken Effekt besitzen pfadbasierte Intentionen des Abbiegens (links- und rechtsabbie-
gen), wohingegen bei einer Geradeausfahrt eine Landmarke selten referenziert wird. Es läßt sich schließen, daß Geradeausfahrten keiner weiteren raumbezogenen Spezifikation bedürfen und als Standardfall angenommen werden. Dies ist dadurch begründet, daß die Aktion des Geradeausfahrens räumlich gesehen eindimensional verläuft und nur eine geringe Winkelvariation zwischen dem aktuellen und dem nächsten Pfadabschnitt zu berücksichtigen ist. Im Gegensatz dazu werden Abbiegeaktionen bezüglich eines zweidimensionalen Raums definiert, wodurch eine genauere Beschreibung notwendig ist, um den intendierten Pfad eindeutig zu referenzieren.

Abb. 3.11 zeigt den Vergleich von relativen Häufigkeiten der Referenzierungen von Höhen- und Farbmerkmalen a) relativ zu Beschreibungen mit Landmarken und b) relativ zu allen Beschreibungen.

Bei einem Abbiegevorgang werden Landmarken fast grundsätzlich verwendet. Besitzt eine referenzierte Landmarke ein auffälliges Farbmerkmal, so wird sie in 73% aller Fälle durch dieses genauer spezifiziert. Bei einem auffälligen Höhenmerkmal erfolgt dies in 93% der Fälle. Die Referenzierung einer Landmarke kommt eine gebietskonstituierende Bedeutung zu. Damit der Rezipient dieses Gebiet eindeutig identifizieren kann, versucht der Sprecher, die Landmarke so genau wie möglich zu beschreiben. Es werden aber grundsätzlich nur solche Landmarken referenziert, die auf der „richtigen“ Seite liegen. In den Fällen, in denen die Probanden Landmarken auf der „falschen“ Seite integrieren sollten, hatten sie erhebliche Probleme. Bei einem Probanden führte dies sogar dazu, daß er nicht in der Lage war, die Aktion rechtzeitig zu beschreiben. Dies ist ein weiteres Indiz für die Schemahäufigkeit der Generierung von Wegbeschreibungen (vgl. [Ke79], [Wu82], [RS87]).

Das Verhältnis zwischen der Verwendung von visuellen Merkmalen Höhe und Farbe bei einem Abbiegevorgang zeigt im Vergleich, daß das Merkmal Höhe auf sprachlicher Ebene von stärkerem Einfluß ist als Farbe, was im Widerspruch zur Erkennbarkeit dieser Merkmale steht (vgl. [Ma86]). Im freien und im festgelegten Beschreibungsfall ergibt sich, daß ein Höhenmerkmal doppelt so häufig sprachlich spezifiziert wird, wie ein Farbmerkmal (vgl. Abbildung 3.11). Dies unterstützt eine Asymmetrievermutung zwischen visueller Wahrnehmung und sprachstruktureller Umsetzung visueller Merkmale (vgl. [Ma86]).
Kapitel 4

Formale Beschreibung des kognitiven egbeschreibungsagenten


Entsprechend dieser Fragestellung wird bezüglich des Phänomens der inkrementellen Wegbeschreibungen ein funktionales Agentenmodell vorgestellt, welches (synthetische) visuelle Daten, Karteninformation, navigationsrelevante Information und Annahmen über den Hörer in zeitliches und räumliches Wissen umsetzt, um das sprachliche Verhalten des Agenten daran anzupassen. Für den Agenten stellen sich in erster Linie Auswahlprobleme. Die Information, die ihm durch eine Situation und eine Karte zur Verfügung steht, ist weitaus umfangreicher als das, was er zum Erfüllen seiner Aufgabe benötigt. Das Auswahlproblem wirft die Frage auf, wie ein


### 4.1 Modularer Aufbau des funktionalen Modells


Das visuelle Modul wählt, basierend auf physikalischen, visuellen Merkmalen aus einer Situation visuell saliente Objekte aus. Das Modul zur Pfadsuche bestimmt durch heuristische, inkrementelle Strategien die zu verfolgenden Pfadabschnitte. Auf der Ausgabeseite übernimmt das Sprachproduktionsmodul die Aufgabe, relevantes Wissen zu versprachlichen und dieses zu ei-
4.1. MODULARER AUFBAU DES FUNKTIONALEN MODELLS

Abbildung 4.1: Module des funktionalen Modells

nem geeignetem Zeitpunkt zu präsentieren. Das Navigationsmodul umfaßt Funktionen, mit denen Bewegungen und damit allgemein die Navigationsfähigkeiten des Agenten auf einer abstrakten Ebene verbunden sind. Es ergibt sich eine vernetzte, funktionale Struktur (siehe Abbildung 4.1). Das Steuerungsmodul koordiniert die Aktivierungen und Deaktivierungen von Modulen und nimmt dadurch eine zentrale Stellung ein. Im folgenden wird zuerst das Steuerungsmodul (Kapitel 4.1.1), dann die Eingabemodule (Kapitel 4.1.2, 4.1.3), das Raumintegrationsmodul (Kapitel 4.1.4) und abschließend das Ausgabemodul (Kapitel 4.1.5, 4.1.6) beschrieben.

4.1.1 Das Steuerungsmodul


Das Steuerungsmodul nimmt eine prominente Stellung im Modell ein. Die einzige Steuerung, die das Steuerungsmoduls ausübt, ist die Aktivation und Deaktivierung von Modulen entsprechend

¹In Fällen, in denen der Agent diese Internalisierung nicht vollführt, kann man umgangssprachlich davon sprechen, daß der Agent „die Zeit vergessen“ hat.

4.1.2 Das Pfadsuchmodul

Verfügt ein Akteur über kein ausreichendes räumliches Langzeitwissen bezüglich einer Umgebung, ist er auf externe Informationsquellen angewiesen. Liegt das Zielobjekt nicht innerhalb des Perzeptionsbereichs, ist es eine Standardkulturtechnik, Straßen- bzw. Umgebungskarten zu verwenden (vgl. [DL88]).


Abbildung 4.2: Transformation eines geometrischen 3D-Modells in eine 2D-Straßenkarten-Repräsentation

Das integrierte Pfadsuchmodul wählt heuristisch und inkrementell aus einer Kartenrepräsentation bezogen auf eine Start- und Ziellokation Pfadschnitte aus. Es wird ein zweistufiger Ansatz verwendet. Auf der Ebene der geometrischen Repräsentation einer Umgebung sind Pfadabschnitte und Verzweigungspunkte definiert. Diese werden auf einer Kartenrepräsentationsebene, die in Form von nichtgerichteten Graphen gegeben ist, abgebildet. Entsprechend den Ergebnissen über Segmentierungsfakten bei Pfaden (siehe Kapitel 2.3.5), besteht eine direkte Abbildung zwischen
4.1. MODULARER AUFBAU DES FUNKTIONALEN MODELLS


---

**Definition 4: Ein Pfadabschnitt**


---

**Definition 5: Ein Verzweigungspunkt**

Ein *Verzweigungspunkt* ist eine Lokation, an dem ein Betrachter bzw. Agent die Orientierung bzw. die Änderung der Orientierung des egozentrischen Referenzsystems verändert.

---

Grundlegend für die Definition von Pfadabschnitten und Verzweigungspunkten ist die Orientierung bzw. die Änderung der Orientierung des egozentrischen Referenzsystems des Agenten. Hierbei spielt die Domäne eine wesentliche Rolle. Die Etablierung eines Verzweigungspunktes ist

---

2Informelle Untersuchungen indizieren, daß zwischen diesen beiden Strategien gewechselt werden kann, bzw. daß die Länge der Strecke, die vorausgeschalt wird, variabel ist.
durch die Festlegung eines Pfadabschnitts determiniert. Ein Verzweigungspunkt wird im einfachsten Falle durch das Ende eines linearen Pfadabschnitts festgelegt (siehe Abbildung 4.3). Jedoch können auch Pfadabschnitte, die Krümmungen enthalten, als Einheit identifiziert werden.\(^3\) Auf einem sich schängelnden Pfad wechselt die Ausrichtung des egozentrischen Referenzsystems permanent. Ist der Agent in der Lage, die globale räumliche Struktur und das Muster der Änderung der Orientierung zu erkennen, so werden diese Krümmungen bzw. die zu verfolgende Trajektorie zu einem Pfadabschnitt zusammengefaßt. Ist er dazu nicht in der Lage, so bleibt nichts anderes übrig, als kleine Pfadabschnitte zu definieren, die von einer Änderung der Orientierung zur nächsten reichen.

![Abbildung 4.3: a) linearer Pfadabschnitt, b) nicht-linearer Pfadabschnitt](image)


### 4.1.3 Das Objektauswahlmodul


\(^3\) Ein dahingehend zu überprüfendes Kriterium wäre das der *Stetigkeit.*
4.1. MODULARER AUFBAU DES FUNKTIONALEN MODELLS

In bisherigen Ansätzen zur sprachlichen Beschreibung visueller Daten, geht man davon aus, daß immer eindeutig vorgegeben ist, welche Objekte in die Beschreibung eingehen sollen (z.B. [NN83b], [ABHR86]). In Arbeiten zur Bestimmung von räumlichen und sprachlich-räumlichen Relationen werden Referenzobjekte und zu lokalisierende Objekte meist explizit angegeben, so daß das Problem der Objektauswahl weitgehend umgangen wird (z.B. [OMT94], [Gap94]). Ein Ansatz zur Auswahl von Referenzobjekten findet sich in [Gap96].


Allgemein ist zu untersuchen, welche Funktionen es erlauben, nur die „richtige“ visuelle Information zu analysieren. Diese Frage ist ad hoc gesehen zirkulär, da ein Betrachter in einer Situation nur dann ein Objekt als salient bezeichnen kann, wenn vorher ein Mindestmaß an Analyse in die gesamte visuell-zugängliche Information investiert worden ist. Eine andere Strategie wäre ein wahlloses Durchforsten der visuellen Information. Es wird in dieser Arbeit ein alternativer Ansatz vorgestellt, in dem eine Situation auf einer Merkmalsbene minimal analysiert wird, um Bereiche zu finden, in denen visuell saliente Referenzobjekte zu finden sind.

Mit der Frage nach der Auswahl von visueller Information assoziiert sind formale Modelle, in denen die visuelle Wahrnehmung als bottom-up oder top-down Verfahren modelliert wird. Erstere sind datengetriebene Modelle. Top-down-Verfahren hingegen sind wissensbasiert in dem Sinne, daß sie den Kontext, die Ziele des Agenten und sein langfristig gespeichertes Wissen berücksichtigen. Das Problem eines reinen bottom-up Verfahrens liegt in der Orientierungsgesichertheit, da es bei der Analyse relevant ist, wozu der Agent die Information benötigt. Andererseits bewirkt eine top-down Analyse, daß in manchen Situationen nach Entitäten und Gesamtzusammenhängen gesucht wird, die in der gegebenen Situation nicht vorhanden sind.


\footnote{Z.B. fällt ein roter Punkt unter vielen blauen auf, wohingegen er unter vielen anderen roten unauffällig ist und somit an Salienz für den Betrachter verliert.}

\footnote{Beispielsweise hat ein Glas voll mit Wasser für einen durstigen Menschen in der Wüste eine andere Bedeutung als in einem Weinlokal.

---

6 Im Straßenverkehr hat die Farbe rot eine funktionale Bedeutung und wird dadurch viel eher fokussiert als in anderen Situationen.
7 Ansätze, die dies nicht berücksichtigen, führen dazu, daß eine gegebene räumliche Anordnung von Objekten in unterschiedlichen Situationen und bei unterschiedlichen Intentionen immer die gleiche partielle Ordnung mit konstanten Salienzwerten liefert (siehe u.a. den Ansatz von [GKS89], [Gar96]).
8 In nicht-städtischer Umgebung übernehmen andere physische Entitäten die Funktion einer Landmarke, wie beispielsweise Flüsse, Bäume oder Felsformationen (vgl. [Sch92]).
9 Rückschlüsse auf pfadgeleitete Intentionen, mittels derer der räumliche Aufmerksamkeitsbereich gesteuert wird, geben die Beschreibungen der empirischen Untersuchungen (siehe Kapitel 3). Besonders interessante Einblicke erlaubten solche Beschreibungen, die unter der Vorgabe gegeben wurden, bestimmte Landmarken an einer Kreuzung zu integrieren. War die Landmarke auf der Abgeseite, so gab die Testperson ohne zu zögern typische Wegbeschreibungen. Lag sie jedoch auf der gegenüberliegenden Seite, hatten die Testpersonen Schwierigkeiten, überhaupt eine Beschreibung zu generieren. Es wäre eine zu grobe Vereinfachung anzunehmen, daß dies auf visuelle Funktionen zurückzuführen ist. Dennoch kann festgestellt werden, daß es für den Beschreiber unwichtig ist, Landmarken von der gegenüberliegenden Seite zu integrieren. Darum ist aus Effizienzgründen anzunehmen, daß Information über Objekte vom Betrachter frühzeitig als aktuell unwichtig ausgefiltert werden, die mit den aktuellen Intentionen nicht konform sind.
4.1.4 Das Raumintegrationsmodul


Abbildung 4.4: Der funktionale Aufbau des Raumintegrationsmoduls

Im weiteren Verlauf werden die verschiedenen Referenzsysteme skizziert, die die Basis für die Etablierung räumlichen Wissens bilden. Anschließend werden die unterschiedlichen Arten von räumlichen Relationen und die Bedeutung des räumlichen Kurzzeitspeichers vorgestellt.

4.1.4.1 Räumliche Referenzsysteme

Räumliche Referenzsysteme lassen sich danach unterscheiden, ob sie egozentrisch, und damit innerhalb, oder allozentrisch außerhalb des Agenten verankert sind (vgl. 2.3.1). Übliche Defi-
nitionen von allozentrischen Referenzsystemen abstrahieren derart vom Betrachterstandpunkt des Agenten, daß sie für einen sich bewegenden Agenten nicht direkt übernommen werden können (vgl. Kapitel 2.3.1). Denn die Eigenbewegung des Agenten bewirkt gleichsam eine dynamische Veränderung der wahrgenommenen, räumlichen Situation. In statischen Situationen (vgl. [Her86], [Gap94]) kann vom Betrachterstandpunkt abstrahiert werden. In dynamischen Situationen ist jedoch die Lokation des Agenten, und damit verbunden vor allem das egozentrische Referenzsystem, im Mittelpunkt der Analyse. Die Orientierung und Verwendung von Referenzsystemen ändert sich dabei dynamisch. Um die Bewegung und die Perspektive getrennt voneinander berücksichtigen zu können, wird das egozentrische Referenzsystem in ein primäres und ein sekundäres Referenzsystem unterteilt. Ein primäres Referenzsystem ist durch die vom Agenten aktuell ausgeführte Hauptidekt der Bewegung, der Handlung oder Perzeption festgelegt. Daneben ist ein Agent durch kurzfristige Umorientierung in der Lage, räumliche Bereiche außerhalb der Hauptidekt zu fokussieren, wozu er ein egozentrisches, sekundäres Referenzsystem, bei gleichzeitigem Erhalt des primären Referenzsystems verwendet. Um Objekte, bezogen auf die Lokation des Agenten, zueinander in Beziehung zu setzen, wird entsprechend dem Prinzip des geringsten kognitiven Aufwands ein geeignetes Referenzsystem ausgewählt. Die Verwendung eines sekundären sowie des nachfolgend beschriebenen virtuellen Referenzsystems kann auch sprachlich explizit angeben werden, wie zum Beispiel durch folgende Anweisung: „Wenn Sie nach rechts schauen, dann ...“


4.1.4.2 Evaluation räumlicher Relationen

Ist bezüglich einer Situation ein geeignetes Referenzsystem ausgewählt worden, werden zwischen ausgewählten Objekten und dem Agenten räumliche Relationen etabliert. Grundlage für Raumbeschreibungen ist eine ausreichende raumbezogene Repräsentation, um geeignete raumsprachliche Ausdrücke zu instantiern. Allgemein repräsentieren verschiedene Typen von Relationen unterschiedliche Arten von konzeptionellen Beziehungen (vgl. Kapitel 2.3.2). In dieser Arbeit werden drei Typen von räumlichen Relationen verwendet, die als Basis der Raumrepräsentationen verwendet werden. Topologische räumliche Relationen repräsentieren Beziehungen, die referenzsystemunabhängig sind. Sie explizieren grobe physische Zusammenhänge auf qualitative Weise. Deiktische und instrinsische räumliche Relationen hingegen hängen vom verwendeten Re-

10Zugrunde liegt die Hypothese, daß der Sprecher annimmt, daß der Rezipient räumliche Referenzsysteme in derselben Weise wie der Sprecher etabliert. Es ist zu vermuten, daß eine solche Strategie erlernt werden muß, und daß sie demzufolge sprachabhängig ist (vgl. [GM95], [GW95]).

11Wie empirische Untersuchungen zeigen, tendieren Probanden dazu, den kognitiven Aufwand des Rezipienten zu minimieren. Da im Falle der inkrementellen Wegbeschreibungen, so wie sie hier verstanden werden, davon ausgegangen wird, daß der Sprecher und der Rezipient sich nebeneinander in die gleiche Richtung bewegen, wird in dieser Arbeit der Aspekt der Einnahme einer Rezipientenperspektive nicht weiter diskutiert. Generell ist eine solche Einschränkung für inkrementelle Wegbeschreibungen nicht notwendig. Jedoch ist eine Beschreibung für einen Rezipienten am einfachsten zu verstehen, wenn sie bezüglich seines egozentrischen Referenzsystems gegeben werden (vgl. [HGH91]).

12Im Sinne von Herrmanns „Sichhineinsetzen“ (vgl. [HGH91]).


Abbildung 4.5: Der konzeptionelle Zusammenhang zwischen den Repräsentationsebenen

Um den Unterschied zwischen der amodal-räumlichen und sprachlichen Ebene zu kennzeichnen, werden räumliche Relationen auf der konzeptuellen Ebene durch englischsprachige Ausdrücke beschildert und durch ”#” prä- und postfigiert, wie zum Beispiel bei #right-of#. Dies kennzeichnet eine explizite Unterscheidung zwischen räumlichen Relationen und räumlichen Präpositionen bzw. deren Lexemen (siehe Abbildung 4.5). Die restriktierte Menge dieser Relationen besteht aus den folgenden Elementen, die allgemein der Menge der projektierten Relationen zugeordnet wird (vgl. [RS88], [Gap94]): #left-of#, #right-of#, in-front-of#, #behind#, #above#
4.1.4.3 Räumlicher Kurzzeitspeicher


4.1.5 Das Navigationsmodul


Es ist prinzipiell notwendig, das raumbezogene Wissen über einen Pfadabschnitt mit Information, die aus der Situation extrahiert worden ist, zu vergleichen. Da es sich im vorliegenden Modell nicht um eine physikalische Realisierung handelt, wird eine dediziertere Instantiierung vernachlässigt. Bei einer Integration des Modells in einen physikalischen Agenten bedarf es einer genauen Formulierung dieser Instantiierung. Eine vollständige Modellierung dieses Moduls fordert die Integration kontrolltheoretischer Ansätze zu physischen Agenten (z.B. [Bro86]), was aber außerhalb des Rahmens dieser Arbeit liegt.

13Es bestehen Ähnlichkeiten des Konzepts der Anwendbarkeitsgrade zu Konzepten der “fuzzy set theory” (z.B. [Zad65], [Lak73]).
14Für einen ersten Ansatz zur Modellierung eines vollständigen Raumintegrationsmoduls siehe [GKSS9].
4.2. Funktionale Beschreibung

4.1.6 Das Sprachproduktionsmodul


4.2 Funktionale Beschreibung des kognitiven Wegbeschreibungsagents

Der Agent agiert in einem dreidimensionalen Modell des Campus der Universität des Saarlandes, welcher mit Landmarken, Pfaden und Verzweigungspunkten drei wesentliche Elemente einer stadtähnlichen Umgebung besitzt (vgl. Abbildung 4.6 und Kapitel 2.3). Mit diesen Elementen korrespondieren Straßen, Kreuzungen, Plätze, Kreisverkehre, Gebäude und Grundflächen. Die Umgebung repräsentiert somit eine Granularitätsstufe, durch die ein Betrachter Zugang zu einer typischen, alltäglichen Umgebung findet.

In einer Situation dieser dreidimensionalen Umgebung selektiert, internalisiert, verarbeitet und verwendet der Agent visuelle Information als Basis für sein sprachliches und navigationsbezogenes Verhalten. Es wird im folgenden das funktionale Modell eines kognitiven Agenten erläutert, welcher in der Lage ist, sich in einer dreidimensionalen Umgebung zu bewegen und adäquate inkrementelle Wegbeschreibungen zu generieren. Grob lassen sich Funktionen unterscheiden, die

\[15\] Nicht modelliert sind bisher dynamische Objekte, wie beispielsweise Autos, Fahrräder oder andere Personen.
in der Schnittstelle zwischen der Umgebung und dem Agenten liegen, und zum anderen solche, die davon abgekapselt und von außen nicht direkt zugänglich sind. Das äußere Verhalten des Agenten besteht in seinen Bewegungen in der Umgebung und seinen verbalen Beschreibungen, die er einem ihm begleitenden Rezipienten vermittelt. Der Agent ist maximal kooperativ und antizipiert in eingeschränktem Maße die Intentionen und das Wissen über den Adressaten, um sein navigationsbezogenes und verbales Verhalten darauf einzustellen.

von raum-zeitlichen Beschränkungen abhängig und bedingt die Verwendung von “anytime”-Mechanismen. 
In den folgenden Kapiteln werden die Funktionen der einzelnen Module beschrieben.

4.2.1 Funktionale Beschreibung des Steuerungsmoduls

Die Aufgabe des Steuerungsmoduls ist die Bestimmung von Ablaufplänen (schedules), durch die Phasen aktiviert und deaktiviert werden. Über Phasen wiederum werden Module angesteuert, die wiederum die Evaluation von Funktionen anstoßen. Das Zusammenspiel von Modulen und Funktionen wird durch zeitliche Restriktionen eingeschränkt, die durch die Eigenbewegung des Agenten und die Umgebung gegeben sind. 
Im weiteren Verlauf wird diskutiert, wie zeitliche Restriktionen bei der Generierung berücksichtigt werden. Dazu wird die Zeit vom Agenten geschätzt, über die er für eine Beschreibung verfügt. Dieser Zeitrahmen wird auf die einzelnen Module verteilt. Anschließend wird vorgestellt, welche Phasen durch das Steuerungsmodul angesteuert werden können und wie diese die einzelnen Verarbeitungsschritte determinieren.
4.2.1.1 Zeitrahmen und Zeitbeschränkungen

Wie die empirischen Untersuchungen zeigen (vgl. Kapitel 3), lassen sich inkrementelle Wegbeschreibungen in fünf Klassen $T_1, ..., T_5$ einteilen, die in direkter Abhängigkeit vom zeitlichen Rahmen der Situation stehen (siehe Tabelle 3.1 in Kapitel 3.2).

Besteht die Beschreibung einer Situation und einer Aktion aus mehreren Teilbeschreibungen, so lassen sich Abhängigkeiten zwischen diesen durch Zuordnung zu Zeitrahmen formalisieren. Der Agent versucht dabei, den Rezipienten nicht mit Information zu überfordern, sondern reduziert die Komplexität auf eine Untermenge von möglichen Beschreibungen. Wird eine Teilbeschreibung zu einem Zeitrahmen $T_i$ gegeben, so wird frühstens zu einem Zeitrahmen $T_{i+2}$ eine weitere Teilbeschreibung präsentiert. Ein solches Verhalten ist nur dann möglich, wenn $i \in \{2, 3\}$ gilt, da während $T_i$ ausschließlich unterstützende Beschreibungen generiert werden (vgl. Kapitel 3.2). Dadurch ergeben sich empirisch belegte Abhängigkeiten, wie sie in Abbildung 4.8 dargestellt sind.

Abbildung 4.8: Zeitintervalle, die den inkrementellen Wegbeschreibungen prototypisch zugeordnet sind.

Um bestimmen zu können, in welchem Zeitrahmen sich der Agent zu einem gegebenen Zeitpunkt befindet, wird der Zeitrahmen $T_{guess}$ vom Agenten geschätzt. Dieser Zeitrahmen determiniert, wieviel Zeit dem Agenten bis zum Erreichen des nächsten Verzweigungspunktes und für die gesamte kommunikative Handlung zur Verfügung steht. Dazu wird $T_{guess}$ in vier Teilzeitrahmen aufgeteilt:

1. $T_{proc}$: Zeitrahmen zur Generierung einer inkrementellen Wegbeschreibung (IWB)
2. $T_{pres}$: Zeitrahmen zur Präsentation einer IWB

---

Finkler schlägt eine Unterscheidung in quantitative und qualitative Inkrementalität vor (vgl. [Fin06]). Ein Prozeß ist demnach quantitativ inkrementell, wenn die Ausgabe in seiner Gesamtheit erst dann zu erkennen ist, wenn alle Einschrittsgaben vollständig gegeben worden sind. Jede Einzelausgabe ist in sich vollständig. Im Gegensatz dazu ist ein Prozeß qualitativ inkrementell, wenn jede Einzelausgabe die Qualität der gesamten Ausgabe verbessert. Ein Beispiel für eine quantitative Inkrementalität ist der zeilenweise Bildaufbau beim Fernseher. Im Gegensatz dazu wird bei einem Mikroskop ein unscharfes Bild qualitativ inkrementell aufgebaut, wenn man die Schärfe verbessert. Dies Unterscheidung wird durch den hier vorgestellten Ansatz „überlappt, da zuerst eine minimale Beschreibung generiert wird, die die Situation vollständig beschreibt. Danach bzw. nebenläufig dazu wird, falls zeitliche Restriktionen dies erlauben, eine Erweiterung dieser Beschreibung generiert, die ihrerseits wiederum vollständig ist. Eine rein qualitative Inkrementalität kann in einfacher Weise dadurch erreicht werden, daß die Erweiterung nicht mehr durch eine eigene vollständige Beschreibung verbalisiert wird, sondern durch Anhängung der neuen Information an die vorhergehende Beschreibung. Da ein solches sprachliches Verhalten im Korpus nicht vorkam, wird dies im folgenden nicht weiter betrachtet.
4.2. FUNKTIONALE BESCHREIBUNG

3. $T_{\text{listener}}$: Angenommener Zeithorizont, den der Empfänger zum Verstehen einer IWB benötigt

4. $T_{\text{action}}$: Geschätzter Zeithorizont zur Ausführungsaktivierung$^{17}$

Die beiden Zeithorizonte $T_{\text{pres}}$ und $T_{\text{listener}}$ hängen voneinander ab, da die Komplexität der Beschreibung approximativ durch die Länge der IWB gegeben ist.$^{18}$ Es wird angenommen, daß sich mit einer längeren Beschreibung auch die Verstehenszeit des Empfängers verlängert. Ein zeitliche Beschränkung, die im Falle der inkrementellen Wegbeschreibungen nicht überschritten werden darf, legt die folgende Ungleichung fest:

$$T_{\text{guess}} \geq T_{\text{proc}} + T_{\text{pres}} + T_{\text{listener}} + T_{\text{action}}$$

Durch Modifikation der internen Einstellungen des Agenten bezüglich den antizipierten Eigenschaften des Empfängers, kann die gesamte Generierung beeinflußt werden. Geht der Agent davon aus, daß der Höhrer eine lange Zeitspanne benötigt, um eine Beschreibung zu verstehen, so bleibt weniger Zeit für die Generierung zur Verfügung, so daß die Qualität der Beschreibung abnimmt. Ergo expliziert $T_{\text{guess}}$ die zeitliche Ressourcenlimitation des Agenten zur Beschreibung des Weges (siehe Kapitel 4.2.1).$^{19}$ Nur wenn diese Beschränkung erfüllt sind, ist sichergestellt, daß die Beschreibung zum richtigen Zeitpunkt erfolgt.$^{20}$

Die Verbindung zwischen den Zeithorizonten $T_i$ ($i = 1 \ldots 5$) und $T_{\text{guess}}$ erfolgt über $T_{\text{pres}}$. Der Agent bestimmt, zu welchem Zeithorizont $T_i$ er intendiert, eine Beschreibung dem Empfänger präsentierbar zur Verfügung zu stellen. Dafür ist die Optimierungsaufgabe zu lösen, die nach dem Zeithorizont $T_i$ mit minimalen $i$ sucht, welcher die untere Grenze der Addition der beiden Zeithorizonte $T_{\text{proc}}$ und $T_{\text{pres}}$ enthält:

$$\text{Suche minimales } i: lb^{21}(T_{\text{proc}} + T_{\text{pres}}) \in T_i$$

Je weniger komplex die Generierung einer Beschreibung ist, um so kürzer sind $T_{\text{proc}}$ und $T_{\text{pres}}$. Da der Agent aber maximal kooperativ ist, sucht er ein $T_i$, welches eine möglichst detaillierte Beschreibung ermöglicht. Die Länge von $T_{\text{proc}}$ und $T_{\text{pres}}$ sind von der Komplexität

$^{17}$Bei Fußgängern kann dieser Zeithorizont vernachlässigt werden. Bei der Navigation von größeren Objekten, wie beispielsweise Öltankschiffen, ist dies jedoch ein ganz entscheidender Faktor, der ein Vielfaches der Zeit einnimmt, die man für die Generierung einer Beschreibung benötigt.

$^{18}$Allgemein ist die Komplexität eines Sates durch die Phrasenstruktur bestimmt (vgl. [Kin77]). Im Falle inkrementeller Wegbeschreibungen hingegen werden durchweg einfache Satzkonstruktionen verwendet, bei denen eine einzelne Phrase auf der Oberflächenstrukturbene fast ausschließlich in maximal drei Worte transformiert wird. Bei kurzen Beschreibungen werden Phrasen meist durch einzelne Worte verbalisiert. Somit erscheint die approximative Abschätzung der Komplexität des Inhalts einer Beschreibung durch die Anzahl der Einzelworte im Falle inkrementeller Wegbeschreibungen gerechtfertigt.

$^{19}$Ein solches Verhalten steht im Einklang mit der Hypothese, daß ein menschlicher Agent über beschränkte Ressourcen verfügt und dabei mit steigender Verfügbarkeit einer Ressource auch die Qualität des Ergebnisses, im vorliegenden Falle des Sprachverhaltens, steigert (vgl. auch [Cha87], [DB87], [RW91]).

$^{20}$Da auch die Probanden nicht in jedem Fall zum richtigen Zeitpunkt eine Beschreibung geben konnten, kann davon ausgegangen werden, daß auch sie nicht immer die Situation zeitlich korrekt einzuschätzen wissen. Dies kann zum einen daher rühren, daß sie die Situation falsch eingeschätzt haben, oder aber bereits von Anfang an wußten, daß sie die Aufgabe nicht erfüllen konnten. Im letzteren Fall waren sie dann nicht in der Lage, die Beschreibung geeignet zu verkürzen (vgl. Kapitel 3).

$^{21}$Die Abkürzung $lb$ steht für die untere Grenze (lower bound).
der Beschreibung abhängig, welche wiederum einem bestimmten $T_i$ zugeordnet ist. Dadurch sind $T_{proc}$ und $T_{pres}$ direkt von $T_i$ abhängig. Somit kann obige Bedingung wie folgt umgeformt werden:

\[
\text{Suche minimales } i: \text{lb}(T_{proc}(T_i) + T_{pres}(T_i)) \geq \text{lb}(T_i).
\]

Aus den Untersuchungen konnten Mittelwerte für $T_{proc}(T_i)$ und $T_{pres}(T_i)$ abgeleitet werden, so daß diese Suche nach einem minimalen $i$ effizient evaluiert werden kann (ein Beispiel hierzu ist in Abbildung 4.9 \(^{22}\) dargestellt).

Im folgenden wird vorgestellt, wie im Modell der Zeitrahmen $T_{proc}$ dekomponiert und dadurch das Verhalten des gesamten Modells bestimmt wird.

### 4.2.1.2 Dekomposition des Zeitrahmens $T_{proc}$

Für das sprachliche Verhalten des Agenten ist in erster Linie der Zeitrahmen $T_{proc}$ verantwortlich. Die Funktionen zur *Evaluation temporaler Beschränkungen* dekomponieren $T_{proc}$ in kleinere Zeitrahmen. Eine statische Dekomposition von $T_{proc}$ auf Module schränkt die Adapтивität des Agenten erheblich ein, was eine geringe Anpassungsfähigkeit zur Folge hat. Adaptiveres Verhalten erlaubt eine Dekomposition in Abhängigkeit vom vorgegebenen Zeitrahmen $T_{proc}$. In diesen, den $T_i$ assoziierten Kombinationen von Inhaltsstrukturen, ist der Zeitrahmen $T_{proc}(T_i)$ \(^{23}\) unterschiedlich lang, was Auswirkungen auf die Qualität der Beschreibungen hat. Die empirisch belegte Einteilung in fünf Zeitrahmen $T_1$ bis $T_5$ erlaubt eine Diskretisierung des Problemraums, so daß vermeiden wird, permanent den Wert von $T_{proc}$ überprüfen und berücksichtigen zu müssen. Wie die empirischen Untersuchungen zeigen (vgl. Kapitel 3), reicht eine verhältnismäßig grobe Schätzung von $T_{proc}$ aus, um eine adäquate Beschreibung zu generieren. Dieser Umstand macht das Verhalten des Agenten robust gegenüber kleinen Änderungen in der Umgebung und gegenüber Fehleinschätzungen.

Unter der Annahme, daß der Sprecher versucht, $T_{proc}$ maximal auszuschöpfen, ergibt sich folgende Ungleichung (vgl. auch Abbildung 4.10):

\(^{22}\)Aus Platzgründen sind die eingezeichneten Zeitrahmen für $T_{proc}$ und $T_{pres}$ nicht in Sekunden skaliert eingzeichnet.

\(^{23}\)In Fällen, in denen es nicht von primärer Bedeutung ist oder aus dem Kontext hervorgeht, bezeugt welchem $T_i$ eine Beschreibung generiert wird, wird die $T_{proc}(T_i)$ durch $T_{proc}$ abgekürzt.
4.2. FUNKTIONALE BESCHREIBUNG

Abbildung 4.10: Kontrollfluß bedingt durch die zeitliche Beschränkung $T_{proc}$

$$T_{proc} \geq T_{object} + T_{ps} + T_{ri} + T_{nlp} + T_{nv}.$$

Die Funktionen der einzelnen Module werden nicht in beliebiger, sondern in partiell koordinierter Weise aktiviert. Zwischen Funktionen bestehen *Sequentialisierungsrelationen*, die angeben, ob zwei Funktionen nur strikt sequentiell oder auch nebeneinander evaluiert werden können.

Durch Abhängigkeiten zwischen Modulen wird angegeben, welche Module zueinander *nebenläufig* beziehungsweise *sequentiell* evaluiert werden können. Ein Modul ist zu einem anderen *nebenläufig*, wenn es einen Zeitpunkt $t$ gibt, an dem beide Module gleichzeitig aktiviert sind. Zwei Module stehen zueinander in sequentieller Beziehung, wenn sie nicht nebeneinander sind.

Zwischen den eingeführten Modulen des Agenten gibt es *Sequentialisierungsrelationen*, die in Tabelle 4.1 abgebildet sind. Hierbei drückt eine Sequentialisierungsabhängigkeit $s(x, y)$, aus, daß ein Modul $x$ nicht gleichzeitig mit Modul $y$ aktiv sein kann. Diese beiden Module können zeitlich voneinander getrennt aktiv sein, hingegen muß zu einem Zeitpunkt mindestens eines von beiden de-aktiviert sein.

Ist beispielsweise das visuelle Objektauswahlmodul VOA-M aktiv, kann das Pfadsuchmodul PS-M nicht aktiv sein und vice versa (vgl. Tabelle 4.1). Dies repräsentiert, daß der Agent nicht gleichzeitig auf eine Karte lesen und die visuelle Information der Situation analysieren kann.

Tabelle 4.1: Sequenzialisierungsabhängigkeiten zwischen Modulen (s: sequentiell, n: nebeneinläufig)

erst eine räumliche Repräsentation bestimmt werden muß, bevor eine Inhaltsstruktur konstruiert
werden kann.²⁴

Sequenzialisierungsrelationen repräsentieren keinen Kontrollfluß, sondern bestimmen nur die Akti-
vationszusammenhänge auf lokaler Ebene zwischen jeweils zwei Modulen. Es ist Aufgabe des Steu-
erungsmoduls unter Beachtung der Sequenzialisierungsabhängigkeiten, Module in geeig-
teter Weise zu aktivieren und zu de-aktivieren. Eine solche Aktivierungsreihenfolge wird durch Phasen
festgelegt, die im folgenden beschrieben werden.

4.2.1.3 Aktivierung und Deaktivierung von Phasen

Eine Phase determiniert die Reihenfolge, in der Module aktiviert werden. Mit jeder Aktivation
ist ein Zeitintervall assoziiert, welches dem Modul angibt, welcher Zeitrahmen für eine bestimmte
Aufgabe zur Verfügung steht. Im vorgeschlagenen Modell werden vier Phasen der Verarbeitung
unterschieden:

1. $P\_schedule$: Erstellung eines Ablaufplans
2. $P\_path$: Suche eines geeigneten Pfades (Etablierung von Routenwissen)
3. $P\_object$: Auswahl von Landmarken nach visuellen Salienzkriterien (Etablierung von Land-
markenwissen)
4. $P\_ndp$: Generierung und Präsentation inkrementeller Wegbeschreibungen

Über den $Typ$ einer Phase wird ein Modul mitgeteilt, in welchem Modus es aktiviert wor-
den ist. Wird das Raumintegrationsmodul über die Phase $P\_path$ aktiviert, werden im RI-M Funktionen anders evaluiert, als wenn die Aktivation durch $P\_object$ erfolgt. D.h., daß durch den
Typ einer Phase Kontrollinformation an die Module weitergegeben wird, die die Evaluation von
Funktionen innerhalb des Moduls beeinflußt. Über die Evaluation der Funktionen wacht alleinig
das jeweilige Modul. Eine derartige Modularisierung der Verarbeitung verringert den Kontroll-
laufwand. Die vier Phasen, die in diesem Modell integriert sind, lassen sich in zwei Klassen
aufteilen. Die erste Klasse enthält die Phase $P\_schedule$, die ausschließlich kontrollflußtechnische
Auswirkungen besitzt. Die zweite Klasse betrifft Phasen, die inhaltsbezogen definiert sind.

Die Phase $P\_schedule$ ist alleinig auf das Steuerungsmodul K-M bezogen. $P\_schedule$ veranlaßt das
Steuerungsmodul einen neuen, an $T_{guess}$ angepaßten, Ablaufplan zu erstellen (vgl. Phase 1).

²⁴Ein anderer Ansatz wäre, beide Module nebeneinläufig zu verwenden, wodurch erreicht wird, daß eine Änderung
in der Konfiguration augenblicklich die Auswahl einer anderen Menge von Inhaltsstrukturen nach sich zieht. Da
aber in den Untersuchungen ein solcher Fall nicht vorkam, wird von einer sequentiellen Abfolge ausgegangen
(s(RI-M, NLP-M)).
4.2. FUNKTIONALE BESCHREIBUNG

\[ P_{\text{schedule}}: \text{Activate-Control-Module}(T_{\text{proc}}); \]

Phase 1: Phase der Ablaufplanbestimmung und -ausführung

Die anderen Phasen beziehen sich auf inhaltliche Aspekte. In der Phase zur Auswahl von Pfadsegmenten \( P_{\text{path}} \) (vgl. Phase 2) wird zuerst die heuristische, inkrementelle Pfadsuche aktiviert, welcher \( T_{ps} \) als Zeitrahmen zur Verfügung hat (Schritt 1 und 2 in Abbildung 4.11). Danach wird das \( PS-M \) deaktiviert und das \( RI-M \) aktiviert (Schritt 3). Die relevanten räumlichen Relationen werden anschließend ausgewertet (Schritte 4 und 5). Das Ergebnis wird im Zeitrahmen \( T_{ri} \) durch das Raumintegrationsmodul in eine sogenannte \textit{räumliche Konfigurationsbeschreibung} integriert (vgl. Phase 2).

\[ P_{\text{path}}: \text{Activate-Path-Selection-Module}(T_{ps}); \]
\[ \text{Activate-Spatial-Module}(T_{ri}); \]

Phase 2: Phase der Pfadsuche und -integration

Abbildung 4.11: Evaluationsschritte der Phase \( P_{\text{path}} \)

Wie die empirischen Untersuchungen zeigen, verzichteten die Testpersonen in restriktiven Situationen auf die Beschreibung von Landmarken und bezogen sich ausschließlich auf Aktionen und Straßenelemente. Ist der Zeitrahmen \( T_{\text{proc}} \) hinreichend groß, so wird durch die Phase \( P_{\text{object}} \) das Modul zur visuellen Objektauswahl evaluiert (vgl. Phase 3), welche referenzbildend eine
Menge von salienten Objekten selektiert (Schritte 1, 2 und 3 in Abbildung 4.12). Die Anzahl der selektrierten Objekte hängt sowohl vom vorgegebenen Zeiträumen $T_{object}$ (vgl. Phase 3), als auch von der Situation selbst ab. Hierdurch werden die Objekte, auf die gleiche Weise wie vorher die Pfadabstraktion, in die räumliche Repräsentationsstruktur durch Evaluation geeigneter räumlicher Relationen integriert (Schritte 4 und 5).

<table>
<thead>
<tr>
<th align="left">$P_{object}$:</th>
<th align="left">Activate-Visual-Module($T_{object}$);</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left"></td>
<td align="left">Activate-Spatial-Module($T_{rl}$);</td>
</tr>
</tbody>
</table>

Phase 3: Phase der visuellen Objektauswahl und -integration

Abbildung 4.12: Evaluationsschritte der Phase $P_{object}$

Das sprachliche Verhalten wird letztendlich über die Sprachgenerierungs- und die Präsentationsphase $F_{nlp}$ gesteuert (vgl. Phase 4). Zuerst werden unter Berücksichtigung von $T_{nlg}$ an die Situation und den Zielen des Agenten angepaßte Inhaltsstrukturen bestimmt (Schritt 1 in Abbildung 4.13), durch Sprachfunktionen in Oberflächenstrukturen (Schritt 2) und nach der Generierung in Verbalstrukturen transformiert (Schritt 4). Anschließend wird das Präsentationsmodul aktiviert, welches darüber wacht, daß die Beschreibung im richtigen Zeitrahmen $T_{nlp}$ gegeben wird (Schritt 5). Im folgenden werden Ablaufpläne beschrieben, die einzelne Phasen zu komplexen Verarbeitungsschritte integrieren.

4.2.1.4 Aktivierung und Evaluation von Ablaufplänen

Einzelne Phasen erlauben noch kein koordiniertes Gesamtverhalten. In einem weiteren Hierarchisierungsschritt werden die Phasen zu Ablaufplänen integriert. Es wird zwischen minimalen und erweiterten Ablaufplänen unterschieden. Sie definieren eine Unterscheidung in reine Aktionenbeschreibungen und solche, die durch Referenzierung von Lokationen und Objekten die
Phase 4: Phase der Sprachgenerierung und Präsentation

Abbildung 4.13: Evaluationsschritte bei der Ausführung von $P_{nlp}$

Beschreibung erweitern.


Beispiel 1: Ablaufplan AP-1

In Beispiel 1 werden in einem ersten Schritt die Phasen $P_I$ und $P_2$ aktiviert. Ein Zeitrahmen wird von einer Phase zerlegt und an die zu aktivierenden Module propagiert. Sobald die Blockierungen von $P_I$ und $P_2$ aufgehoben sind, werden die Phasen der nächsten Teilliste aktiviert, was im obigen Beispiel $P_3$ ist. Es werden in einem Ablaufplan dann zwei Phasen aktiviert, wenn die zweite Phase vom Typ $P_{schedule}$ ist, wodurch ein alternativer Ablaufplan zum aktuellen bestimmt wird. Die Aktivierung von $P_{schedule}$ bewirkt keine Blockierung nachfolgender Phasen. Der aktuelle aktive Ablaufplan ist durch die globale Variable *schedule* zugreifbar.

Ein minimaler Ablaufplan gewährleistet, daß der Agent sich an neue Situationen grundlegend
anpaßt. Daraus resultierende Beschreibungen enthalten die minimale Information, welche sichert, daß der Rezipient eine, wenn auch minimale, so doch adäquate Beschreibung erhält. Bei einem minimalen Ablaufplan (vgl. Beispiel 2) werden zuerst geeignete Pfadabschnitte gesucht und zueinander in Beziehung gesetzt \( P_{\text{path}} \), \( t_{\text{path}} \). Nebenläufig dazu wird ein erweiterter Ablaufplan \( (P_{\text{schedule}}, t_{\text{schedule}}) \) generiert, wie in Abbildung 4.14 illustriert ist.

\[
A_{\text{P_{minimal}}} = (((P_{\text{path}}, t_{\text{path}})(P_{\text{schedule}}, t_{\text{schedule}})))
\]

**Beispiel 2: Minimaler Ablaufplan**

Diese Verzweigungsmöglichkeit des Planungsansatzes beruht auf der Unabhängigkeit der einzelnen Module. Ist ein erweiterter Ablaufplan bestimmt worden, so wird dieser in der gleichen Weise wie der minimale abgearbeitet. Für den Fall, daß eine bereits aktive Phase erneut aktiviert werden soll, muß damit solange gewartet werden, bis diese Phase durch den vorhergehenden Ablaufplan deaktiviert wurde. Dadurch werden Überschneidungen und Inkonistenzen vermieden.

\[
A_{\text{P_{erweitert}}} = (((P_{\text{path}}, t_{\text{path}})(P_{\text{schedule}}, t_{\text{schedule}})))
\]

**Abbildung 4.14: Nebenläufige Ausführung von minimalen und erweiterten Ablaufplänen**

Ein mögliche Erweiterung des Ablaufplans besteht in der Integration der Objektauswahl-Phase \( (P_{\text{object}}, t_{\text{object}}) \). Da während \( A_{\text{P_{minimal}}} \) die aktuellen Pfadabschnitte bereits ausgewählt worden sind, ist es nicht notwendig \( P_{\text{path}} \) erneut zu aktivieren (vgl. Beispiel 3).

\[
A_{\text{P_{extended}}} = (((P_{\text{object}}, t_{\text{object}})(P_{\text{schedule}}, t_{\text{schedule}})))
\]

**Beispiel 3: Erweiterter Ablaufplan**

Die Integration einer Landmarke in die räumliche Konfigurationsbeschreibung \( (P_{\text{object}}, t_{\text{object}}) \) hat zur Folge, daß eine um eine Ortsreferenz erweiterte Inhaltsstruktur selektiert und in eine natürlichsprachliche Beschreibung umgesetzt wird. Wiederum wird nebenauftrag zu \( P_{\text{object}} \) erneut \( P_{\text{schedule}} \) aktiviert, um zu überprüfen, ob weitere Landmarken integriert und damit die Beschreibung verbessert werden kann (vgl. Abbildung 4.14). Dies hängt einerseits davon ab, ob es in der Situation weitere saliente Landmarken gibt, und davon, ob noch ein genügend großer Zeitrahmen \( T_{\text{guess}} \) zur Verfügung steht.

Abschließend werden die Funktionen des Steuerungsmoduls vorgestellt, die die Steuerung des gesamten Modells innerhalb.
4.2. FUNKTIONALE BESCHREIBUNG

4.2.1.5 Funktionen des Steuerungsmoduls

Das Verhalten des Steuerungsmoduls wird durch eine globale Schleife determiniert, die initial einen minimalen Ablaufplan bestimmt und überprüft, ob noch auszuführende Ablaufpläne vorhanden sind. Ist dies nicht der Fall, so befindet sich der Agent in einem Endzustand. Die Ausführung von Ablaufplänen erfolgt durch die Funktion *Evaluate-Schedule* (vgl. Funktion 1). Diese Funktion startet die Evaluation zur Bestimmung zeitlicher Beschränkungen (*Evaluate-Time-Frame*), die in der Funktion *Determine-Schedule* dazu verwendet werden, einen adäquaten Ablaufplan zu bestimmen (siehe Funktion 2). Die Ausführung eines Ablaufplans erfolgt durch die Funktion *Execute-Schedule* (siehe Abbildung 4.15).

Abbildung 4.15: Aktivationszyklus zwischen den Funktionen der globalen Koordination und der Evaluation zeitlicher Beschränkungen im Steuerungsmodul

Initialisiert wird die globale Funktion *General-Control-Loop* durch einen minimalen Ablaufplan, der durch die Funktion *Determine-Minimal-Schedule* generiert wird (vgl. Funktion 2). Dieser Plan ist so konzipiert, daß der Agent ausschließlich eine minimale, auf Routenwissen basierende Beschreibung generiert. Ist der zur Verfügung stehende Zeithorizont \( T_{\text{guess}} \) kleiner als der vom minimalen Ablaufplan geforderte, so kann keine Beschreibung generiert werden. Der Zeithorizont \( T_{\text{guess}} \) wird durch die Funktion *Evaluate-Time-Frame* bestimmt. Dazu bestimmt *Estimate-Time-Frame* die Strecke zwischen der aktuellen Lokation des Agenten (\( \text{loc(ego)} \))\(^{25}\) und der Lokation des nächsten Verzweigungspunktes (\( \text{loc(ep)} \)). Mit Hilfe der aktuellen Geschwindigkeit wird unter Verwendung der Gleichung \( t = s/v \) geschätzt, wie lange der Agent benötigt, den Verzweigungspunkt zu erreichen. Entsprechend \( T_{\text{guess}} \) wird über die Funktion *Determine-Schedule* ein geeigneter Ablaufplan ausgewählt (vgl. Funktion 1).

Entsprechend diesem Ablaufplan werden Phasen aktiviert und deaktiviert. Die Kontrollschleife wird verlassen, wenn *schedule* die Nachricht *finished* enthält. Ist es nicht möglich, zu \( T_{\text{guess}} \) einen alternativen Plan zu erzeugen, geht die Funktion *Execute-Schedule* vorübergehend in einen Wartezustand, bis die Beschreibung präsentiert worden ist, die sich noch im Präsentationsspeicher befindet. Nach Ausführung der Aktion befindet sich der Agent in einer neuen Situation, was zur Folge hat, daß der minimale Ablaufplan erneut aktiviert wird (vgl. Funktion 1).

\(^{25}\)Dies wird der Einfachheit wegen durch *ego* abgekürzt.
Funktion 1: Evaluate-Schedule

\[\text{Evaluate-Schedule (schedule)}\]
\[
\text{if (situation is not new) then} \\
\text{if (for all } s \text{ in schedule: for all } p \text{ in phases(s): (In-Time(p))) then} \\
\text{Execute-Schedule(s);} \\
\text{else} \\
\text{schedule := Determine-Schedule(Estimate-Time-Frame(loc(ego), loc(ep(s))));} \\
\text{Execute-Schedule(schedule);} \\
\text{else} \\
\text{Initialize-Schedule; } \\
\text{schedule := Determine-Minimal-Schedule(s);} \\
\text{Execute-Schedule(schedule);}
\]

In konstanten zeitlichen Abständen wird überprüft, ob die einzelnen Phasen die ihnen zugewiesenen Zeiträume überschritten haben. Bei einer Überschreitung wird der zugehörige Ablaufplan aus der Liste *schedule* entfernt. Falls der Agent sich jedoch in einer neuen Situation befindet, wird *schedule* initialisiert und ein minimaler Ablaufplan aktiviert (vgl. Funktion 1).

Funktion 2: General-Control-Loop

\[\text{General-Control-Loop}\]
\[
\text{schedule := Determine-Minimal-Schedule(Estimate-Time-Slot(s));} \\
\text{while (finished is not in "schedule") do} \\
\text{Evaluate-Schedule("schedule");} \\
\text{done}
\]

Sollte eine der Phasen den ihr zugeteilten Zeitrahmen überschritten haben, wird ein alternativer Ablaufplan erstellt. Im einfachsten Fall heißt dies, daß \( T_{guess} \) zu gering ist, um einen alternativen Ablaufplan zu bestimmen. In diesem Falle kann der Agent nur die Beschreibung präsentieren, die vom letzten vollständig durchlaufenen Ablaufplan generiert worden ist. Ist \( T_{guess} \) ausreichend groß, so wird ein alternativer Ablaufplan erstellt. Die aktuellen Ablaufpläne werden durch die Funktion *Execute-Schedule* ausgeführt bzw. kontrolliert (vgl. Funktion 3).

Ist das erste Element eines Ablaufplans \( s \) abgearbeitet worden, werden die Phasen des restlichen Ablaufplans aktiviert (vgl. Funktion 3). Befindet sich die erste zu aktivierende Phase in einem Wartezustand, so verbliebt der gesamte Ablaufplan solange im Wartestatus, bis die jeweilige Blockierung aufgehoben ist. Diese Planungsvorgehensweise ließe sich in mancher Hinsicht verbessern. Beispielsweise könnte der Zeitrahmen einer Phase dynamisch modifiziert werden. Folglich müßte ein komplexer Planungsmechanismus integriert werden, welcher evaluiert, ob und wie Ergebnisse bisheriger Berechnungen weiterverwendet werden können. Kann der Teilplan nicht weiter verwendet werden, wäre ebenso eine einfachere Strategie erfolgreich, in der eine Phase sofort deaktiviert wird, sobald ein alternativer Aufruf dieser Phase geplant worden ist. Weiterhin wäre möglich, daß das Steuerungsmodul an der inhaltlichen Planung in den Modulen beteiligt wäre. Dies wiederum würde zu Interferenzen zwischen Phasen und Modulen führen, die nur durch komplexe Schlußfolgerungsmechanismen berechnet werden können. Vor diesem Hintergrund ist der verwendete Planungsansatz einfach, besitzt aber die Eigenschaft, Kontrollwissen und inhaltbezogenes Wissen strikt zu trennen, was den Planungsaufwand minimiert und die Planungsstrategie effizient einsetzbar macht.
Funktion 3: Execute-Schedule

\[\text{Execute-Schedule}(sched)\]
\text{for all } s \text{ in } sched; \\
\text{case:} \\
\text{(Deactive-Status (first(s))):} \\
\hspace{1em} s := \text{rest}(s); \\
\text{Activate-Phases(first (s));} \\
\text{(Waiting-Status (first(s))):} \\
\hspace{1em} \text{if (not Already-Activated-Phases(first(s))) then} \\
\hspace{2em} \text{Activate-Phases(first (s));} \\
\hspace{1em} \text{if (sched is not nil) and (sched is not in *schedule*) then} \\
\hspace{2em} \text{*schedule* := (union *schedule* sched);} \\
\hspace{1em} \text{if (not Already-Activated-Phases(first(s))) then} \\
\hspace{2em} \text{Activate-Phases(first schedule);} \\
\text{else} \\
\hspace{1em} \text{(Set-Into-Waiting-Status(first schedule));} \\
\]

Im weiteren werden die inhaltssorientierten Module vorgestellt. Dabei wird zur einfacheren Beschreibung von zeitlichen Beschränkungen und der Kommunikation mit dem Steuerungsmodul abstrahiert.

4.2.2 Funktionale Beschreibung des Pfadsuchmoduls


Als Grundlage für die Beschreibung eines Weges muß der Agent wissen, welche Sequenz von Pfadabschnitten zu verfolgen ist, um die Ziellokation zu erreichen. Die inkrementelle, heuristische Pfadsuche erfolgt in drei Schritten. Zuerst wird mittels der Pfadsuchfunktion \text{Incremental-Heuristic-Path-Search}, der auf den aktuellen Pfadabschnitt \((P_{\text{pred}(vp)} \rightarrow vp)\) folgende Pfadabschnitt \((P_{vp} \rightarrow \text{succ}(vp))\) determiniert.26

Als Bedingung an eine Sequenz von Pfadabschnitten gilt, daß es in dieser Sequenz zu jedem Pfadabschnitt \(P_a\) einen weiteren Pfadabschnitt \(P_b\) gibt, der direkt angrenzt. Das wird durch eine topologische Relation \(#\text{meets#}(P_a P_b)\) repräsentiert. Dies indiziert, daß es eine Lokation \(VP\) gibt, mit \(P_{x \rightarrow vp} = P_a\) und \(P_{vp \rightarrow y} = P_b\) die den Verzweigungspunkt zwischen diesen beiden Pfadabschnitten definiert (siehe Abbildung ??).


Im folgenden werden die Funktionen der inkrementellen heuristischen Pfadsuche im einzelnen vorgestellt.

---

26Dabei steht \(vp\) für Verzweigungspunkt, \(\text{succ}(vp)\) ist der \(vp\) nachfolgende und \(\text{pred}(vp)\) der vorhergehende Verzweigungspunkt.
4.2.2.1 Funktionen der inkrementellen heuristischen Pfadsuche

Zur Berechnung eines geeigneten Nachfolgepfadschnitts bestimmt die inkrementelle, heuristische Suchfunktion einen kostenminimalen Nachfolgebogen (move) zum Endbogen (state) der aktuellen Kante (P_{pred(state)} → state) (vgl. Abbildung 4.17). Die Kosten setzen sich dabei aus den tatsächlichen Kosten zusammen, die benötigt werden, um den Knoten move zu erreichen und denen, die man von move aus entsprechend der Schätzung zum Zielenknoten benötigt. Ist ein Nachfolgebogen bestimmt, so wird davon ein Pfadschnitt (P_{state} → move) abgeleitet. Da davon ausgegangen wird, daß der Agent sich in einer für ihn neuen Umgebung befindet, wird der Pfad inkrementell mit einer begrenzten Vorausschau (look-ahead) bestimmt.


Abbildung 4.17: Korridor bezüglich des Gebiets um den Startknoten \textit{start} und den Zielknoten \textit{goal}

Die Funktion \textit{Incremental-Heuristic-Path-Search} wird über \( P_{path} \) iterativ aufgerufen, solange der Zielknoten \textit{goal} noch nicht erreicht worden ist. Über den Parameter \textit{limit} wird angegeben, wieviele Schritte im voraus geplant werden, d.h. wieviel sukzessive Knoten und Kanten des Pfades zu einem Zeitpunkt bestimmt werden. Da kein Rückgriff auf raumbezogene Langzeitrepräsentationen verwendet wird, beträgt die Begrenzung \textit{limit} ungefähr eins bis zwei, da dies einer Approximation der Pfadebene entspricht, die über den visuellen Wahrnehmungsbe- reich erschließbar sind.

Der aktuelle Knoten \textit{state} wird expandiert, indem alle direkten Nachfolgeknoten bis zu einer Tiefe von \textit{limit} evaluiert werden (vgl. Funktion 5). Das Ergebnis dieser Evaluation sind die Kosten \( f' \), die dem Nachfolgeknoten zugeordnet sind (vgl. Funktion 4). Alle berechneten Kosten werden in einer Hashtabelle \textit{Table} gespeichert. Der Nachfolgeknoten von \textit{state} ist ein solcher, welchem der geringste Kostenwert \( f' \) zugeordnet ist.

Die zentrale Funktion der Pfadsuche heißt \textit{Evaluation}, in welcher die Kosten eines direkt Nachfolgeknotens berechnet werden (vgl. Funktion 5). In dieser Funktion wird eine Schleife solange durchlaufen, wie die Liste der aktiven Knoten \textit{(Open)} nicht leer ist. Der minimale Kostenwert des Knotens \textit{move} wird in der Variablen \( \alpha \) gespeichert. Entsprechend der im A*-Algorithmus verwendeten Funktion \( f \), setzen sich die Kosten aus den \textit{tatsächlichen} Kosten für die Strecke von einem Knoten \textit{state} zum Nachfolgeknoten \textit{move} und den \textit{heuristisch} geschätzten Kosten für das Erreichen des Zielknotens von \textit{move} zusammen. Ausgehend von \textit{move} werden dessen Nachfolgeknoten daraufhin überprüft, welchem bis zu einer Tiefe \textit{limit} oder dem Erreichen des Zielknotens die geringsten Übergangskosten zugeordnet sind. Knoten, die den aktuellen Wert von \( \alpha \) unterschreiten, werden weiterverfolgt. Als Anpassung an die Domäne physikalischer Umgebungen sind die Übergangskosten \textit{(move-cost)} zwischen einem Knoten \textit{node} und einem Nachfolgeknoten...
Funktion 4: Incremental-Heuristic-Path-Search

\[\text{Incremental-Heuristic-Path-Search}(\text{state}, \text{goal})\]

\[
\text{while (state \neq \text{goal}) do} \\
\quad \text{expand state}; \\
\quad \text{for each child move of state do} \\
\quad \quad f'(\text{move}) := \text{evaluate(move, limit)}; \\
\quad \quad \text{done}; \\
\quad \text{Table(state) = update(state, children)}; \\
\quad \text{state = move with lowest } f' \\
\text{done}
\]


Funktion 5: Evaluation

\[
\text{Evaluation}(\text{move}, \text{limit}) \\
\quad \text{Open := \{move\};} \\
\quad \alpha := \infty; \\
\quad f(\text{move}) := g(\text{move}) + h(\text{move}); \\
\quad \text{while Open not nil do} \\
\quad \quad \text{node = pop Open;} \\
\quad \quad \text{expand node;} \\
\quad \quad \text{for all child of node do} \\
\quad \quad \quad g(\text{child}) = g(\text{node}) + \text{move-cost(node, child)}; \\
\quad \quad \quad f(\text{child}) = g(\text{child}) + h(\text{child, goal}); \\
\quad \quad \quad \text{if } f(\text{child}) < \alpha \text{ then} \\
\quad \quad \quad \quad \text{if depth = limit or goal(child) then} \\
\quad \quad \quad \quad \quad \alpha = f(\text{child}); \\
\quad \quad \quad \quad \text{else} \\
\quad \quad \quad \quad \quad \quad \text{push child on Open;} \\
\quad \quad \quad \text{done} \\
\quad \text{done} \\
\quad \text{return(\alpha)}
\]


\textsuperscript{27}Hierdurch eröffnet sich ein breites Feld für die Diskussion der Verbindung eines Suchalgorithmus und seiner Anwendungsdomäne. Dies soll hier nicht weiter diskutiert werden, da dies zum eigenen Problem der Wegbeschreibung nicht sonderlich viel beiträgt. Angemerkt sei aber, daß diese Dominänenabhängigkeit in der Implementation dieses Modells für verschiedene Bewegungsmodi integriert worden sind (vgl. [AK96]).

\[
\text{Funktion 6: Heuristische Funktion } h \\
\text{dist} = \text{distance} (\text{node}, \text{goal}) / \text{corridor} (\text{node}); \\
\text{return} (\text{dist}).
\]

Die Speicherung von Kostenwerten der Knoten dient einerseits der Effizienzsteigerung und andererseits zur Auffrischung von Zyklen. Hierzu werden die Kosten, die einem Knoten zugeordnet sind, bei jedem nochmaligen Erreichen dieses Knotens um einen Wert $\epsilon$ erhöht. Um Seiteneffekte zu vermeiden, ist der Wert von $\epsilon$ größer Null, aber kleiner gleich den minimalen Kosten zwischen beliebigen zwei Knoten des Suchraums.

\[
\text{Funktion 7: Update} \\
\text{Update} (\text{state}, \text{children}) \\
\text{if} \ \text{Table} (\text{state}) = \text{nil} \ \text{then} \\
\text{return} (\text{min} (f' (\text{children}) ) + \epsilon) \\
\text{else} \\
\text{return} (\text{Table} (\text{state}) + \epsilon)
\]

### 4.2.2.2 Pfadabschnitte und Verzweigungspunkte

Gleichzeitig mit der Auswahl des kostengünstigsten Nachfolgknotens move wird die Kante $P_{\text{state} \rightarrow \text{move}}$ festgelegt, über die man diesen Knoten vom Endknoten state der aktuell aktiven Kante aus erreichen kann. Diese Kante wird auf einen in der Situation identifizierbaren Pfadabschnitt $P_{\text{state}} \rightarrow \text{move}$ abgebildet. Ist ein Pfadabschnitt $P_{\text{start}} \rightarrow \text{end}$ identifiziert worden, so wird er auf verschiedene Weise repräsentiert (Abstract-Path-Segment). Wie in späteren Verarbeitungsschritten deutlich wird, kann hierdurch eine Verarbeitung von pfadbezogenem Wissen erreicht werden. Ein Pfadabschnitt ist zum einen durch seinen Schwerpunkt (cog($P$); center of gravity) und andererseits durch seinen Anfangs- (start) und Endpunkt (end) repräsentiert.\footnote{Durch die zu start und end assoziierten Lokationen ist die Richtung des Pfadabschnitts $P_A \rightarrow B$ vorgegeben.} Ein Verzweigungspunkt $VF$ wird zusätzlich durch alle Knoten, die mit den Anfangsknoten von Wege übereinstimmen und seinem Schwerpunkt repräsentiert. Im letzten Verarbeitungsschritt der Phase $\text{P}_\text{path}$ werden Pfadabschnitte in räumliche Konfigurationsbeschreibungen integriert (vgl. Abbildung 4.16). Hierzu werden deiktische und intrinsicische räumliche Relationen zwischen dem neuen Pfadabschnitt und dem Agenten zu bereits integrierten Pfadsegmenten und zu selektierten Landmarken, die räumlich nahe sind, etabliert (vgl. Kapitel 4.2.4).
Abbildung 4.18: Ableitung eines punktförmigen Verzweigungspunkts a) ohne und b) mit alternativen Pfadabschnitten

Im weiteren wird vorgestellt, Verzweigungspunkte, und dabei insbesondere freie Plätze und Kreisverkehre, behandelt werden.

4.2.2.3 Expansion von Verzweigungspunkten

Aus der Definition eines Pfadabschnitts (vgl. Definition 4) geht hervor, daß ein Übergang von einem Pfadabschnitt zum nächsten dann stattfindet, wenn sich die Ausrichtung des egozentrischen Referenzsystems qualitativ ändert. Eine Änderung dieses Referenzsystems bedingt in der Domäne der Wegbeschreibungen eine Rotation, insbesondere bezüglich der Frontalachse, in der horizontalen Ebene. Befindet sich der Agent an einer Lokation, an der eine solche Rotation stattfindet, können alternative Pfadabschnitten vorhanden sein. Auf einem Verzweigungspunkt verläßt der Agent die eindimensionale Struktur des Pfadabschnitts und beritt einen zweidimensionalen Raum, in dem er sich in zwei Dimension bewegen kann. Ein Verzweigungspunkt selbst, soweit es kein Platz oder Kreisverkehr ist, wird als punktförmig angenommen. Die Bestimmung der Knoten move und succ(move) etabliert gleichzeitig Verzweigungspunkte VP(move) und VP(succ(move)) (siehe Abbildung 4.18). Für die Pfadabschnitte P_state \rightarrow move und P_move \rightarrow succ(move) müssen die topologischen Relationen gelten, damit sie begehbar sind (\#meets\# P_state \rightarrow move P_move \rightarrow succ(move)), mit VP_move \in P_state \rightarrow move \cap P_move \rightarrow succ(move).


Für eine Beschreibung dieser Pfadabschnitte reicht diese Repräsentation nicht aus. Es wird im weiteren davon ausgegangen, daß zwischen den Pfadabschnitten P_state \rightarrow move und P_move \rightarrow succ(move) ein freier Platz (VP_move) liegt, bezüglich dem auf Grund sei-
4.2. FUNKTIONALE BESCHREIBUNG

Abbildung 4.19: Expansion eines freien Platzes. Eine punktförmige Repräsentation aus der Kartenrepräsentation (a)) wird in eine expandierte überführt (b))

In der physikalischen Ausdehnung ein eigener Pfadabschnitt spezifiziert werden muß, der es dem Hörer erlaubt, den anschließenden Pfadabschnitt \( P_{\text{move}} \rightarrow \text{succ}(\text{move}) \) zu erreichen. Dazu wird die Fläche, die durch einen freien Platz gegeben wird, in weitere Verzweigungspunkte und Pfadabschnitte dekomponiert. Der Eintrittsort in den freien Platz \( VP_{\text{move}} \) ist gleichzeitig der Endpunkt des Pfadabschnitts \( P_{\text{state}} \rightarrow \text{move} \). Der Knoten \( VP_{\text{move}} \) wird durch einen Expansionsschritt in eine Knoten-Kante-Knoten-Sequenz transformiert (\( P_{\text{state}} \rightarrow \text{move} P_{\text{move}} \rightarrow \text{succ}(\text{move}) \)) in (\( P_{\text{state}} \rightarrow \text{move} P_{\text{move},1} \rightarrow \text{move,v2}, P_{\text{move}} \rightarrow \text{succ}(\text{move}) \)). Die beiden Verzweigungspunkte \( VP_{\text{move},1} \) und \( VP_{\text{move},2} \) werden durch einen virtuellen Pfadabschnitt \( P_{\text{move},1} \rightarrow \text{move,v2} \) verbunden, der keine direkte Entsprechung in der Knoten- bzw. Kantenrepräsentation der Karte besitzt (vgl. Abbildung 4.19b).


\[
\begin{align*}
(P_{\text{state}} \rightarrow \text{move} P_{\text{move}} \rightarrow \text{succ}(\text{move})) \\
\Rightarrow \quad (P_{\text{state}} \rightarrow \text{vp,v1} P_{\text{vp,v1}} \rightarrow \text{vp,v2} P_{\text{vp,v2}} \rightarrow \text{vp,v3} \\
P_{\text{vp,v3}} \rightarrow \text{vp,v4} P_{\text{vp,v4}} \rightarrow \text{VP succ(move)})
\end{align*}
\]

Beispiel 4: Beispiel des expandierten Ablaufplans eines Kreisverkehrs

Ergibt die Pfadsuche folgende Abfolge der Pfadelemente (\( P_{\text{state}} \rightarrow \text{move} P_{\text{move}} \rightarrow \text{succ}(\text{move}) \))
4.2.2.4 Zusammenfassung

Der hier vorgestellte inkrementelle, heuristische Suchalgorithmus basiert auf dem unidirek
4.2. FUNKTIONALE BESCHREIBUNG

Neben Information über Pfadabschnitte werden in inkrementellen Wegbeschreibungen visuell perzipierte Objekte referenziert. Geeignete Funktionen zur Auswahl von Objekten werden im folgenden Kapitel diskutiert.

4.2.3 Funktionen des Objektauswahlmoduls


4.2.3.1 Allgemeine Beschreibung der Auswahl visuell salienter Objekte


\(^{31}\) Ob die Information über die visuellen Merkmale eines Objektes durch visuelle Wahrnehmungsprozesse oder durch einen Zugriff auf ein digitalisiertes Objektmodell erhalten wird, ist in diesem Ansatz von sekundärem Interesse.
Funktion 8: Select-Salient-Object

\[ \text{Select-Salient-Object}(S, \text{intention}, T_{\text{object}}) \]

Determine-Visible-Objects:
\[ \text{sit} := \text{Project-Situation}(\text{Situation}); \]
Determine-Focus($S, \text{intention}, T_{\text{object}}$):
raster-rep := Discretize-Into-Raster(sit);
\[ \text{f-maps} := \text{Decompose-Into-Feature-Maps}(\text{raster-rep}); \]
\[ \text{s-maps} := \text{Determine-Salience-In-F-Maps}(\text{f-maps}); \]
\[ \text{i-map} := \text{Integrate-Feature-Maps}(\text{s-maps}); \]
\[ \text{salient-locs} := \text{Determine-Salient-Locs}(\text{i-map}); \]
\[ \text{salient-objects} := \text{Determine-Salient-Objects}(\text{salient-locs}); \]

Das Ergebnis der Funktion \text{Decompose-Into-Feature-Maps} ist die Dekomposition in typspezifische Merkmalskarten (f-maps). Auf jeder dieser Merkmalskarten werden saliente Lokationen bestimmt, die angeben, ob ein bestimmtes Merkmal im Verhältnis zur gesamten Situation salient ist (\text{Determine-Salience-In-F-Maps}). Unterschieden wird die Bestimmung von visuellen Salienzen rekursiver und linearer Merkmale sowie die lokationserhaltende Integration dieser in der \text{globalen Merkmalskarte}. Die merkmalspezifische Berechnung visueller Salienzen rekursiver Merkmale ist in Kapitel 4.2.3.6 für das Merkmal „Farbe“ und linearer Merkmale in Kapitel 4.2.3.7 für die Merkmale „Höhe“ und „Breite“ explizit angegeben. Auf Grund der Rekursionseigenschaft werden visuelle Salienzen rekursiver Merkmale auf einer Rasterebene und solche für lineare Merkmale auf Objektebene berechnet. Das Resultat ist die Zuordnung eines Salienzwertes einer Zelle relativ zu einer Merkmalskarte. Überlagert man die einzelnen typspezifischen Merkmalskarten und integriert die Salienzwerte einer jeden Zelle in topographieerhaltender Weise (\text{Integrate-Feature-Maps}), ergibt sich eine merkmalsübergreifende Repräsentation (i-map) salenter Zellen (\text{Determine-Salient-Locs}). Dieser Verarbeitungsschritt befindet sich immer noch auf der Ebene der Rasterrepräsentation. Erst im darauffolgenden Schritt wird die Verbindung zur Objektebene hergestellt, wenn die Referenz von Objekten zu Salienzen einzelner Zellen etabliert wird (\text{Determine-Salient-Objects}). Das Ergebnis der Funktion \text{Select-Salient-Landmarks} sind solche Objekte einer Situation, deren Salienzwert über einem ressourcenaabhängigen Schwellwert liegen. Insgesamt ergibt sich ein funktionaler Zusammenhang, wie er in Abbildung 4.22 illustriert ist. Im weiteren sollen die Funktionen im einzelnen genauer diskutiert werden.

4.2.3.2 Identifikation sichtbarer Objekte

Die allgemeine Bestimmung von visuell-sichtbaren Objekten ist eine komplexe Problemstellung, die hier nur soweit diskutiert werden soll, wie sie für diese Arbeit relevant ist. Das Ergebnis einer vollständigen Identifikation von Objekten ist der Aufbau von multiplen, unterschiedlich abstrahierten Objektrepräsentationen (vgl. [Her96]). Auf der untersten Stufe der Komplexität ist die Schwerpunktrepräsentation, die allein aus einem Punkt bestimmt ist. Daneben sind u.a. zweidimensionale Repräsentationen in der horizontalen Ebene und minimal-umschreibende Quader definiert. Auf einer weitaus komplexeren Ebene befinden sich Repräsentationen von Haupt- und Nebenachsen des Objektes sowie vollständige geometrische Modellbeschreibungen (z.B. [Mar82]; [Bie90]).
Abbildung 4.22: Der funktionale Aufbau des visuellen Objektauswahlmoduls

4.2.3.3 Abbildung der perspektivischen Sicht auf die Projektionsebene


4.2.3.4 Berücksichtigung der räumlichen Fokussierung

Bezüglich dieser Projektionsebene fokussiert der Agent auf Teilbereiche einer Situation. Eng mit dem Konzept eines Fokus ist das Gebiet der selektiven Aufmerksamkeit verbunden. Wie zahlreiche empirische Erhebungen zeigen, ermöglicht eine Beschränkung der Aufmerksamkeit, daß der Zusammenhang zwischen kognitiven Verarbeitungsaufwand und Komplexität der Situation...
nicht linear ist. Vielmehr erlauben selektive Aufmerksamkeitsprozesse, daß selbst in komplexen Situationen noch adäquate Entscheidungen getroffen werden können.\footnote{Der räumliche Fokus findet in der Pfadscheibe im Konzept des \textit{Korsdorfs} eine Entsprechung (vgl. Kapitel 4.2.2).}

\begin{center}
Abbildung 4.23: Parallelprojektion der Situation auf die Projektionsebene
\end{center}

\begin{center}
\begin{footnotesize}
\begin{enumerate}
\item \textit{Determine-Focus (S, intention, T}_{object})
\item \textit{Determine-Center-Of-Focus(intention)};
\item \textit{Determine-Area-Of-Focus(T}_{object});
\end{enumerate}
\end{footnotesize}
\end{center}

\begin{center}
Abbildung 4.24: Ausrichtung des räumlichen Fokus entsprechend pfadbezogener Intention, nach links abbiegen zu wollen, und Berücksichtigung des Öffnungswinkels $\alpha$ und Abweichungswinkels $\beta$.
\end{center}

\[
(1) \quad F_x = K_x \cdot T_{\text{object}}; \quad F_y = K_y \cdot T_{\text{object}}
\]

Abbildung 4.25: a) Perspektivische Sicht, b) Varibabler räumlicher Fokusbereich mit Breite $2 \cdot F_x$.

Unter Annahme eines linearen Zusammenhangs sind die Koeffizienten $K_x$ und $K_y$ an Hand von informellen Befragungen bestimmt worden. $F_x$ und $F_y$ determinieren den Bereich, in dem alle Objekte höchste Aufmerksamkeit besitzen.

### 4.2.3.5 Abbildung der Projektions- auf die Lokationsebene


Der Fehler, der durch die Approximation einer Fläche $A$ bedingt ist, setzt sich bezüglich der

\(^3\) informell ist eine solche Vorgehensweise von Toda in seinem “Fungus Eater” vorgeschlagen worden (vgl. [Toda2]).

4.2. FUNKTIONALE BESCHREIBUNG

Abbildung 4.26: Transformation einer Situationsprojektion in eine diskrete Lokationsrepräsentation

Abbildung 4.27: Größenfehler, die durch die Approximation der Fläche durch das minimal umschließenende Rastergebiet entstehen.

x-Achse $f_x$ und y-Achse $f_y$ aus den Distanzen zur nächsten Zelle zusammen (vgl. Gleichungen 2 und Abbildung 4.27).

\[
\begin{align*}
(2) & & f_x &= \delta_1 + \delta_2 \\
& & f_y &= \delta_3 + \delta_4
\end{align*}
\]

Der Fehler ist im schlechtesten Falle in jeder Dimension kleiner als die Ausdehnung einer einzelnen Zelle $d$ (vgl. Ungleichung 3).

\[
(3) \quad f_x < d; \ f_y < d
\]


\[
(4) \quad f = 2^*(#x \cdot f_x) + (#y \cdot f_y) < 2^* (#x + #y) \cdot d^2
\]

Daraus ergibt sich ein relativer Fehler $\delta f$ (vgl. Ungleichung 5). Der relative Fehler $\delta f$ wird entsprechend den gegebenen zeitlichen Beschränkungen gewählt. Dazu gilt, daß die Größe des
Fehlers in einem umgekehrt-proportionalem Verhältnis zum Berechnungsaufwand steht.

\[
\delta f = \frac{f}{1 + x + y} \leq d^2
\]

Als Ergebnis liefert Funktion Discretize-Into-Raster eine approximierte Rasterrepräsentation, in der jeder Zelle ein Objekt auf der Projektionsebene zugeordnet ist (vgl. Abbildung 4.21).

### 4.2.3.6 Bestimmung visueller Salienzen bezüglich des rekursiven Merkmals „Farbe“

In welcher Weise visuelle Salienzen rekursiver Merkmale berechnet werden können, wird am Beispiel des Merkmals Farbe erläutert. Das rekursive Merkmal Farbe ist für die Etablierung visueller Salienzen von besonderer Bedeutung, da es durch empirische Studien als dominantes Merkmal bei der Auswahl von Objekten belegt ist (vgl. [Man86]).

Es gibt verschiedene Farbwahrnehmungsmodelle, die in unterschiedlicher Weise approximieren, wie der Mensch ein Farbmerkmal wahrnimmt. Beispiele für solche Modelle sind das Munsell System (vgl. [WS82]), das „natural color system“ (NCS) (vgl. [Här66]), das DIN System (vgl. [DIN80]) und das OSA/UCS System (vgl. [Der91]). Eine typische Zerlegung einer Farbe erfolgt, entsprechend der Heringschen Farbentheorie, in Farbwert, Helligkeit und Sättigung. Problematisch ist an diesen Zerlegungen, daß es selbst für das Konzept des Farbwertes keine einheitliche, formal klar spezifizierte Definition gibt (vgl. [Sv96]). Weiterhin sind die resultierenden Farbwahrnehmungsmodelle derart komplex, daß sie für einen Realzeiteinsatz nicht zu verwenden sind. Um festzustellen in wieweit sich ein Farbmerkmal in einer Situation abhebt, muß ihm zunächst ein numerischer Farbwert zugeordnet werden. Ein Farbmerkmal läßt sich in verschiedener Weise eindeutig zergliedern.


Bezogen auf das CIE-Chromatiken-Diagramm, bestimmte McAdam in diesen Untersuchungen elliptische Bereiche um 25 Referenzpunkte, in denen die Betrachter keinen Farbunterschied feststellen konnten (siehe Abbildung 4.28, [WS82, S. 521]).

Ellipsen von McAdam repräsentieren \((S_r = 1,7; S_g = 6,0; S_b = 1)\).\(^{36}\) Durch diese Anpassung des RGB-Modells an Farbwahrnehmungsphänomene beim Menschen, ist in approximativer Weise ein effizientes, kognitiv adäquates Berechnungsmodell für die Bestimmung der Auffälligkeit von Farbmerkmalen möglich. Der Vorteil der Verwendung des angepassten RGB-Modells ist, daß jedes computerbasierte Farbmodell auf das RGB-Modell zurückgreift.\(^{37}\)

Abbildung 4.29: Veränderung von Salienzegenschaften eines Merkmals in verschiedenen Situationen

\(^{36}\)Bei unterschiedlichen Gesamtlichtverhältnissen unterliegt die Wahrnehmung von Farbe jedoch starken Schwankungen, sodaß Erkenntnisse, die durch Laborexperimente bei standardisierten Randbedingungen gewonnen worden sind, nicht ohne weiteres in Realumgebungen übertragen werden können. Phänomene wie der Farbkonstanzeffekt werden in diesem Ansatz nicht berücksichtigt.

4.2.3.6.1 Die Funktionen zur Bestimmung visueller Salienzen des Merkmals „Farbe“

Um feststellen zu können, ob ein Farbmerkmal einer Zelle als salient zu bezeichnen ist, muß definiert werden, wie sich Farbunterschiede feststellen lassen. Zur Ermittlung eines Farbunterschieds ist es notwendig, einen farbbezogenen Referenzvektor bezüglich einer Situation festzulegen. Ist ein Referenzvektor bestimmt, wird der Farbwert der jeweiligen Zelle dazu in Beziehung gesetzt.


\[
(6) \\
C_{ref}(S) = \frac{1}{(N + M)} \left( \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \text{red}(c_{ij})}{\sum_{i=1}^{N} \sum_{j=1}^{M} \text{green}(c_{ij})} \right) \\
\left( \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \text{blue}(c_{ij})}{\sum_{i=1}^{N} \sum_{j=1}^{M}} \right)
\]

Der relative Farbanteil color\((c_{ij})\) eines Farbwertes wird durch den Abstand zum Referenzvektor berechnet. Der resultierende Differenzvektor \( \Delta\text{color}(c_{ij}) \) ist an die Verzerrungen der Farbwahrnehmung angepaßt und repräsentiert den Unterschied zwischen dem Referenzwert \( C_{ref}(S) \) und dem Farbvektor einer Zelle \( \text{color}(c_{ij}) \) (vgl. Gleichung 7).

\[
(7) \\
\Delta\text{color}(c_{ij}) = |C_{ref}(S) - \text{color}(c_{ij})| = \left( \frac{|\text{red}(C_{ref}(S)) - \text{red}(\text{color}(c_{ij}))|}{S_{\text{red}}} \right) \\
\left( \frac{|\text{green}(C_{ref}(S)) - \text{green}(\text{color}(c_{ij}))|}{S_{\text{green}}} \right) \\
\left( \frac{|\text{blue}(C_{ref}(S)) - \text{blue}(\text{color}(c_{ij}))|}{S_{\text{blue}}} \right)
\]

Der Abstand zwischen zwei Farbwerten hängt von der Verteilung in der gesamten Situation ab. In einer Situation, in der eine Vielzahl von unterschiedlichen Farbwerten sichtbar sind, ist das Merkmal Farbe unbedeutender, als in einer solchen mit geringer Farbvariation (vgl. beispielsweise Abbildung 4.29a,c). Repräsentiert wird dieser Sachverhalt durch die Bestimmung der lokalen Verdichtung eines Farbmerkmals einer Zelle \( c_{ij} \) (vgl. Gleichung 8). Die Verdichtung

---

\(^38\) In Abbildung 4.29 ist der rote Farbpunkt der Deutlichkeit halber, durch einen weißen Kreis mit schwarzen Rand dargestellt.

\(^39\) \(c_{ij}\) ist der Farbvektor, der einer Zelle an der Rasterposition \((i,j)\) des Rasters zugeordnet ist; \(\text{red}(x)\), \(\text{green}(x)\) und \(\text{blue}(x)\) sind die Zugriffsfunktionen auf die Komponenten des Farbvektors bei RGB-Zerlegung. In den hier vorgestellten Ansatz werden konstante Gesamtlichtverhältnisse angenommen.
ist ein Maß für die Konzentration eines Farbwertes bezüglich einer Situation.

\[
\text{lokdenst}(\text{color}(c_{ij})) = \frac{1}{\Delta \text{red}(\text{color}(c_{ij})) + \Delta \text{green}(\text{color}(c_{ij})) + \Delta \text{blue}(\text{color}(c_{ij}))} \left( \begin{array}{l}
\text{red}(\Delta \text{color}(c_{ij})) \\
\text{green}(\Delta \text{color}(c_{ij})) \\
\text{blue}(\Delta \text{color}(c_{ij}))
\end{array} \right)
\]

Die einzelnen Farbkomponenten werden nicht unabhängig voneinander verarbeitet, sondern werden in \(\text{lokdenst}(\text{color}(c_{ij}))\) integriert. Bei monochromatischen Farbwerten besitzt genau eine Komponente den Maximalwert 1 und ist somit identisch mit einer RGB-Repräsentation. Im Falle eines gleichen Anteils aller Komponenten, was dem Farbeindruck \(\text{grau}\) entspricht, ist \(\text{lokdenst}(\text{color}(c_{ij}))\) gleich dem Vektor \(\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)\) und indiziert, daß keiner Farbkomponente ein besonderes Gewicht zukommt. Je stärker der Farbwert sich einer monochromatischen Farbe annähert, um so größer wird der Wert in einer Komponente.


\[
\text{globdenst}(\text{color}(S)) = \frac{1}{N \times M} \left( \begin{array}{l}
\sum_{i=1}^{N} \sum_{j=1}^{M} \text{lokdenst}(\text{red}(c_{ij})) \\
\sum_{i=1}^{N} \sum_{j=1}^{M} \text{lokdenst}(\text{green}(c_{ij})) \\
\sum_{i=1}^{N} \sum_{j=1}^{M} \text{lokdenst}(\text{blue}(c_{ij}))
\end{array} \right)
\]

Der Unterschied zu einer einfachen Mittelwertbildung ist, daß \(\text{globdenst}(\text{color}(S))\) die lokalen Verdichtungen und farbwahrnehmungsbedingte Verzerrungen berücksichtigt. Die globale Verdichtung repräsentiert den Farbeindruck der gesamten Situation. Durch die Bestimmung der

\(^{40}\) Eine Weiterführung des hier vorgestellten Ansatzes ist die Bestimmung von gebietsbezogenen Verdichtungen. Eine interessante, empirische Fragestellung ist, welche „cluster“ von Testpersonen erkannt werden und wie sie den Farbkreis in Gebiete einteilen.
globalen Verdichtung einer Situation können die Farbmerkmale einzelner Zellen verglichen werden. Um die visuelle Salienz bezüglich des Merkmals **Farbe** einer Zelle zu bestimmen, wird der relative Farbanteil \( \Delta \text{color}(c_{ij}) \) mit der globalen Verdichtung \( \text{globdens} (\text{color}(S)) \) in Beziehung gesetzt, woraus der Salienzvektor \( \text{sal} (\text{color}(c_{ij})) \) hervorgeht (vgl. Gleichung 10). Dazu wird der relative Farbvektor der Lokation mit der *globalen Verdünnung* \( \epsilon - \text{density} (\text{color}(S)) \) multipliziert.\(^{41}\) Anerkennt, daß sowohl die lokale als auch die globale Verdichtung bezüglich normierten RGB-Werten definiert sind. Dadurch wird ein Vergleich von beiden zur Bestimmung der visuellen Salienz des Farbmerkmals einer Zelle möglich.

\[
\text{sal} (\text{color}(c_{ij})) = \begin{pmatrix}
\Delta (\text{red} (\text{color}(c_{ij})) \times (1 - \text{red} (\text{globdens}(\text{color}(S)))) \\
\Delta (\text{green} (\text{color}(c_{ij})) \times (1 - \text{green} (\text{globdens}(\text{color}(S)))) \\
\Delta (\text{blue} (\text{color}(c_{ij})) \times (1 - \text{blue} (\text{globdens}(\text{color}(S))))
\end{pmatrix}
\]

Durch Betragsbildung ergibt sich der skalare, visuelle Salienzwert des Farbewertes einer Zelle \( \text{sal} (\text{color}(c_{ij})) \) (vgl. Gleichung 11).

\[
\text{sal} (\text{color}(c_{ij})) = \left| \text{sal} (\text{color}(c_{ij})) \right| / s,
\]

mit \( s = \max \left| \text{sal} (\text{color}(c_{ij})) \right| \)

[Bestimmung des salientesten Farbmerkmals][color-map-serra-sgi]

Visuelle Salienzen von Zellen werden an assoziierte Objektepräsentationen auf der Projektionsebene zurückgeführt (vgl. Abbildung 4.21). Dies erfolgt unter der Annahme, daß in der vorliegenden Domäne die Objekte homogene Farbverteilungen besitzen.\(^{42}\) Somit ist der Farbwert eines Objektes gleich dem einer ihm zugeordneten Zellen \( c_{ij} \), sodaß die Berechnung von Salienzen auf

\(^{41}\) Die *globale Verdünnung* ist das Inverse zur *globalen Verdichtung*.

\(^{42}\) Ist ein Objekt nicht homogen gefärbt, wird der Mittelwert über alle assoziierten Zellen bestimmt. Dieses
einer lokationsgebundenen Ebene eine Bedeutung auf Objektebene erhält. Die eher holistische Betrachtungsweise auf Rasterebene besitzt dadurch Einfluß auf die atomistische Betrachtungsweise der Objektebene.

4.2.3.7 Bestimmung visueller Salienzen der linearen Merkmale „Höhe“ und „Breite“

Bei der Verarbeitung von linearen Merkmalen, wie Höhe und Breite, werden Salienzen nicht auf einer Rasterebene berechnet, da lineare Merkmale direkt an Objekte gebunden sind. Als Repräsentationsgrundlage für Höhe und Breite werden minimal umschreibende Rechtecke der Objekte auf der Projektionsebene verwendet, welche die maximale Ausdehnung eines Objektes bezüglich der horizontalen und vertikalen Dimension repräsentieren (vgl. den unteren Teil der Abbildung 4.26). Im weiteren wird, basierend auf der Invarianzannahme, davon ausgegangen, daß die Distanz und somit die Höhe und die Breite eines Objektes vom Agenten hinreichend genau abgeschätzt werden kann.

Für die Bestimmung visueller Salienzen von Höhen- und Breitenmerkmalen, wird nicht auf die absolute Höhe und Breite eines Objektes, sondern vielmehr auf den sichtbaren Anteil auf der Projektionsebene Bezug genommen. Im Gegensatz zu rekursiven Merkmalen lassen sich Höhe und Breite durch jeweils eine einzige Dimension und demzufolge skalär repräsentieren. Der Referenzwert eines linearen Merkmals in einer Situation ergibt sich aus dem arithmetischen Mittel (vgl. die Gleichungen 12, 13).

\[ \bar{H} = \frac{1}{n} \sum_{i=1}^{n} \text{height}(O_i) \]

\[ \bar{W} = \frac{1}{n} \sum_{i=1}^{n} \text{width}(O_i) \]

Die Skalare height(O) und width(O) repräsentieren die Werte der Höhe und der Breite des korrespondierenden Rechtecks eines Objektes O. Um zu vermeiden, daß Objekte, die eine zu geringe Höhe und Breite besitzen, berücksichtigt werden, werden als Erweiterung der “feature integration theory” Schwellwerte eingeführt. Als Randbedingung für die Schwellwerte th\_height und th\_width gilt, daß alle Merkmalswerte größer als der typspezifische Schwellwert sein müssen (vgl. Ungleichung 14).

\[ \text{height}(O_i) > \text{th}_{\text{height}} \land \text{height}(O_i) > \text{th}_{\text{width}} \]

Durch die Invarianzannahme gestützt, wird für die Bestimmung des Salienzwertes bezüglich der Höhen- und Breitenmerkmale eine lineare Metrik verwendet. Die visuellen Salienzen bezüglich des Merkmalstyps Höhe und Breite bilden sich aus der positiven Distanz zwischen dem Merkmalswert (height(O) bzw. width(O)) und dem Referenzwert (H bzw. W) (vgl. die Gleichungen 15).

15, 16).

\[ (15) \quad \text{salement of height}(O) = |\hat{H} - \text{height}(O)| \]

\[ (16) \quad \text{salent of width}(O) = |\hat{W} - \text{width}(O)| \]

Abbildung 4.30: Bestimmung des salientesten Höhenmerkmals


4.2.3.8 Integration in globale Merkmalskarten

Abbildung 4.31: Bestimmung des salientesten Breitenmerkmals

Kapitel 2). Allgemein wird die Frage nach der visuellen Salienz, \( \text{vis-sal} \), eines Objektes durch Gleichung 17 expliziert.

\[
\text{vis-sal}(O) = \alpha \odot \text{sal(color(O))} \odot \beta \odot \text{sal(height(O))} \odot \gamma \odot \text{sal(width(O))}
\]

(17)

Abbildung 4.32: Bestimmung des salientesten Objektes über die globale Merkmalskarte
Sowohl die Koeffizienten der einzelnen Salienzwerte, als auch die Operationen ⊗ und ⊘, sind bisher nicht genauer empirisch untersucht worden. Mangel Evidenzen wird hier eine Linearkombination verwendet. Dazu sind die Koeffizienten empirisch bezüglich der simulierten 3D-Umgebung bestimmt worden (vgl. Kapitel 3). Es ergab sich die Koeffizienten $\alpha = 0.5$, $\beta = 0.3$ und $\gamma = 0.2$.

\[(18)\quad \text{vis-sal}(O) = \alpha \cdot \text{sal}(\text{color}(O)) + \beta \cdot \text{sal}(\text{height}(O)) + \gamma \cdot \text{sal}(\text{width}(O))\]


\[\text{vis-sal}(O) = \alpha \cdot \text{sal}(\text{color}(O)) + \beta \cdot \text{sal}(\text{height}(O)) + \gamma \cdot \text{sal}(\text{width}(O))\]

Abbildung 4.33: Beispiel der Auswahl eines salienten Objektes bei zentraler Fokussierung

4.2.3.9 Adaption der visuellen Salienzen an den räumlichen Fokusbereich


\[(19)\quad \text{vs}(O) = \text{focus}(\text{vis} - \text{sal}(O))\]

\[\text{vis-sal}(O) = \alpha \cdot \text{sal}(\text{color}(O)) + \beta \cdot \text{sal}(\text{height}(O)) + \gamma \cdot \text{sal}(\text{width}(O))\]

Ein interessanter Ansatzpunkt ist es, Koeffizienten für verschiedene Domänen zu bestimmen und miteinander zu vergleichen.
4.2. FUNKTIONALE BESCHREIBUNG

4.2.3.10 Zusammenfassung


Im folgenden Abschnitt wird vorgestellt, wie die durch die Pfadauswahl und die Objektauswahl selezierte Information in räumliches Wissen transformiert wird.

4.2.4 Funktionen des RauminTEGRATIONSmoduls


4.2.4.1 Egozentrische und allozentrische räumliche Referenzsysteme

In statischen Situationen ruht sowohl die Orientierung und Lokation des Betrachters, als auch die der Objekte. Die meisten Arbeiten im Kontext der Raumkognition nehmen solche statischen Situationen an (vgl. Kapitel 2). In diesem Fall hat der Betrachter die Möglichkeit, Objekte durch Lagebeziehungen auf sein statisches Egozentrum zu beziehen oder aber auf eine mental eingenommene Lokation unter Einnahme einer Orientierung, die durch Eigenschaften der Lokation
oder des Objektes *intrinsisch* oder durch den Akteur *extrinsisch* vorgegeben sind (vgl. Kapitel 2.3.1, 2.5). Bewegt sich der Betrachter relativ zu seiner Umgebung, ändern sich einerseits seine Lagebeziehungen zu Objekten und andererseits die räumliche Verankerung der Referenzsysteme und deren Orientierung dynamisch. Hierzu wird die Verwendung von Referenzsystemen auf dynamische Situationen erweitert.


Abbildung 4.34: Bewegung eines Agenten auf einem Pfad in einer Umgebung mit zwei Landmarken A und B zu den Zeitpunkten $t_1$, $t_2$ und $t_3$. 

KAPITEL 4. FORMALE BESCHREIBUNG DES AGENTEN
aktionszeiten bei egozentrierten Referenzsystemen nach einem Anstieg mit Maximum bei 120° wieder ab. Hieraus folgt grundlegend, daß Rotationen von Referenzsystemen beim „Sichhineinversetzen“ mit kognitivem Mehraufwand verbunden sind. Aufgrund kognitiver Ökonomie folgt ebenfalls eine Beweizung egozentrischer Referenzsysteme, die sich auch empirisch nachweisen läßt (vgl. Kapitel 2.3.1, 2.5). Im Elastizitätsbereich hingegen, kann die Orientierung egozentrischer Referenzsysteme nahezu beliebig mental rotiert werden.


Es werden im folgenden drei Referenzsystemklassen (primär, sekundär und virtuell) diskutiert. Dabei wird gezeigt, wie sie dazu verwendet werden, raumbezogenes Wissen in dynamischen Umgebungen zu etablieren.

4.2.4.1.1 Das primäre Referenzsystem Das primäre Referenzsystem (PRS(\(x, y, z\))) basiert auf dem Konzept eines kartesischen Koordinatensystems mit drei unabhängigen Dimensionen, die den dreidimensionalen Raum aufspannen (vgl. Kapitel 2.3.1). Die Ausrichtung der z-Achse (oben/unten) ist parallel zur Gravitationskraft. Die Orientierung der Frontalrichung des primären Referenzsystems ist ausschließlich durch die Hauptschäf festgelegt, die durch die Richtung der Bewegung, der Perzeption oder der Aktivität des Agenten determiniert ist (vgl. Abbildung 4.35). Die Ausrichtung in der Transversalen leitet sich von der Frontalrichtung ab. Bewegt sich der Agent auf einem Pfad, so ist die Hauptschäf durch die Trajektorie der Bewegung determiniert (vgl. Abbildung 4.36). Die Hauptperzeptionsrichtung ist im Falle einer Vorwärtsbewegung mit der Frontalrichtung meist identisch, so daß sich das primäre Referenzsystem wesentlich von der Hauptschäf der Bewegung der Perzeption abkoppelt.

In dynamischen Situationen, in denen der Betrachter über verschiedene Freiheitsgrade verfügt, kann die perzeptive Hauptrichtung kurzfristig modifiziert werden, indem der Agent seine Aufmerksamkeit auf ein bestimmtes Gebiet im Sichtfeld fokussiert. Ein solcher Fall liegt außerhalb der Eigenschaften primärer Referenzsysteme, weswegen es einer zusätzlichen Klasse von Referenzsystemen bedarf.

4.2.4.1.2 Das sekundäre Referenzsystem Bei kurzfristigen Änderungen der Hauptschäf, besonders der Perzeption, verfügt der Agent über ein sekundäres Referenzsystem (SRS), welches es ihm ermöglicht, räumliche Sachverhalte lokal räumlich zu erfassen. Nachdem dieser räumliche Sachverhalt nicht weiter fokussiert wird, wird das sekundäre Referenzsystem
Abbildung 4.35: Orientierung des primären Referenzsystems in der horizontalen Ebene erfolgt immer in Ausrichtung der Hauptachse (entlang Winkel $\alpha = 0^\circ$)

Abbildung 4.36: Primäres Referenzsystem zum Zeitpunkt $t$. Der Agent bewegt sich entlang der Hauptachse des Pfades mit einer Geschwindigkeit $\dot{v}$ an einer räumlichen Konfiguration mit den Objekten $A, C_1, C_2$ und $C_3$ vorbei


\begin{equation}
SRS(x, y, z) = \text{ROT}_\alpha(\text{PRS}(x, y, z))
\end{equation}

Ist ein $SRS$ etabliert, so wird dieses bevorzugt verwendet. Deiktische sprachliche Beschreibungen erfolgen in solchen Fällen ausschließlich bezüglich dieses $SRS$. In Abbildung 4.38 bewegt sich der Agent beispielsweise entlang eines Pfades und perziert zum Zeitpunkt $t_1$ einen räumlichen Sachverhalt, der durch die Landmarken $A, C_1, C_2$ und $C_3$ gegeben ist. Es wird angenommen, daß die Ausrichtung der Aufmerksamkeit in Richtung des Schwerpunkts des räumlichen Sachverhalts durch Rotation um den Winkel $\alpha$ erfolgt. Entsprechend diesem Rotationswinkel wird ein $SRS$
zum Zeitpunkt $t_2$ etabliert. Bewegt sich der Agent mit einer Geschwindigkeit von $\vec{v}$ entlang des Pfades, so erreicht er zu einem Zeitpunkt $t_3$ eine Lokation, bei der der Rotationswinkel $\alpha$ nahezu 90° beträgt. Zum Zeitpunkt $t_1$ sind PRS und SRS in ihrer Orientierung nahezu identisch. Diese Übereinstimmung verstärkt die Verwendung dieser beiden Referenzsysteme zur Beschreibung der Situation (vgl. [CRI93]). Aus diesem Grund wird die Situation zum Zeitpunkt $t_1$ durch Verwendung des sekundären Referenzsystems beschrieben („$C_3$ ist vor A.“). Zum Zeitpunkt $t_2$ ändert sich der Zusammenhang zwischen SRS und PRS. Beide verstärken sich nicht mehr, so daß zwischen beiden gewählt werden muß. Um in dieser Situation $C_3$ noch visuell wahrnehmen zu können, muß SRS eingenommen werden. Von PRS aus ist $C_3$ zum Zeitpunkt $t_2$ nicht mehr zugreifbar, so daß eine entsprechende Beschreibung unangebracht ist (vgl. Satz 21). Somit wird in dieser Situation das SRS ausgewählt.

\[
(21) \quad \text{PRS } t_1: \text{ „}C_3\text{ ist vor A.“} \\
\text{PRS } t_2: \\
\text{SRS } t_1: \text{ „}C_3\text{ ist vor A.“} \\
\text{SRS } t_2: \text{ „}C_3\text{ ist rechts von A.“}
\]


Wie extrinsische und intrinsische Beschreibungen indizieren, ist ein Sprecher in der Lage, auch räumliche Beziehungen relativ zu einem Referenzsystem außerhalb seines Egozentrum zu etablieren. Welche Referenzsysteme hierzu notwendig sind, wird im folgenden vorgestellt.

4.2.4.1.3 \textbf{Das virtuelle Referenzsystem} Allozentrische Referenzsysteme gehören nicht zur Klasse \textit{primärer} und \textit{sekundärer Referenzsysteme}, weswegen eine dritte Klasse von Referenzsystemen erforderlich ist, die \textit{virtuell} in einer beliebigen Lokation einer Situation verankert

\footnote{\textbf{44}Nicht betrachtet wird hierbei die Möglichkeit, den Rezipienten explizit darauf hinzuweisen, ein sekundäres Referenzsystem zu verwenden.}
Abbildung 4.38: Sekundäres Referenzsystems zum Zeitpunkt $t$. Ausrichtung des SRS auf den Schwerpunkt der Konfiguration zu den Zeitpunkten $t_1$ und $t_2$ bei einer Bewegung mit Geschwindigkeit $\vec{v}$.

Abbildung 4.39: Änderung der Hauptachse des $PRS$ zu einem Zeitpunkt $t_2$.

sind (VRS). Im Gegensatz zu einigen raumsprachlichen Arbeiten (vg1. [Cla73], [Her86] und Kapitel 2.5, 2.3.1) wird davon ausgegangen, daß die Etablierung von allozentrischen Referenzsystemen mit unterschiedlichem kognitiven Aufwand verbunden ist. Besonders deutlich wird im Experiment von Herrmann et al. (vg1. Kapitel 2.3.1 und [HGH91]) die Abhängigkeit von Rotationswinkeln beim „Sichhineinversetzen“ in eine Lokation der Situation. Entsprechend dieser Experimentalreihe lassen sich Beschränkungen an virtuellen Referenzsystemen (VRS) definieren. Durch imaginierte Translation wird ein VRS in einer Lokation der Situation außerhalb des Agenten verankert. Es werden solche Rotationswinkel bevorzugt, die sich im Elastizitätsbereich, d.h. einer Winkelabweichung $\alpha$ von $60^\circ$ relativ zur Hauptsache, befinden. Dabei wird angenommen, daß das VRS durch eine imaginierte Rotation aus dem $PRS$ bzw., falls etabliert, aus dem $SRS$, entsteht (vg1. Abbildung 4.40). Formal entspricht dies einer Rotation des $PRS$ bzw. $SRS$ um den Winkel $\alpha$: $\text{ROT}_\alpha$.

\[
\text{VRS}(x, y, z) = \text{ROT}_\alpha(\text{TRANS}_\delta(\text{SRS}(x, y, z)))
\]

mit \(300^\circ \leq \alpha \leq 60^\circ\)

Ein VRS ist, ebenso wie das SRS, lokal bezüglich einer Situation und einer Aufmerksamkeitsausrichtung definiert (siehe Abbildung 4.41). Dabei werden solche VRS bevorzugt, die parallel zu einer der Hauptachsen ($h_1$ und $h_2$ in Abbildung 4.41) des fokussierten räumlichen Sachverhalts liegen (vg1. $VRS_1$ und $VRS_2$ in Abbildung 4.41).
4.2. **Funktionale Beschreibung**

Zwei Fälle sind bei der Etablierung eines virtualen Referenzsystems zu unterscheiden.

1. **Extrinsische** Etablierung eines virtuellen Referenzsystems
2. **Intrinsische** Etablierung eines virtuellen Referenzsystems

Im ersten Fall wird das Referenzsystem außerhalb und im zweiten Fall innerhalb eines Objekts verankert. Die Behandlung einer intrinsischen Etablierung berücksichtigt extrinsische Eigenschaften des Referenzobjekts. Dieser Fall wird in Abschnitt 4.2.4.1.6 diskutiert. Die extrinsische Etablierung eines virtuellen Referenzsystems wird im folgenden betrachtet (vgl. Abschnitt 4.2.4.1.4, 4.2.4.1.5).

**4.2.4.1.4 Etablierung von virtuellen Referenzsystemen bei orthogonalen Konfigurationen** Bei einer extrinsischen Etablierung wird das virtuelle Referenzsystem in einer Lage unmittelbar in der Nähe des fokussierten räumlichen Sachverhalts verankert. In Abbildung 4.41 befindet sich das virtuelle Referenzsystem \( VRS_1(t) \) zum Zeitpunkt \( t \) bezüglich der Hauptachse \( h_1 \) an Lokoation \( l_1 \) bei einer Rotation \( \beta \) relativ zur Orientierung des sekundären Referenzsystems \( SRS(t) \). Gleichzeitig kann ein \( VRS_2(t) \) bezüglich Hauptachse \( h_2 \) an Lokoation \( l_2 \) etabliert werden, was jedoch aufgrund des großen Winkels \( \alpha \) eine geringe Anwendbarkeit besitzt.

**4.2.4.1.5 Etablierung von virtuellen Referenzsystemen bei nicht-orthogonalen Konfigurationen** Bisher ist von Situationen ausgegangen worden, deren Lage sich parallel oder orthogonal zu der Bewegungsrichtung des Agenten fand. Ist dies nicht der Fall, so stehen die Hauptachse des Agenten und die Hauptachsen, \( h_1 \) und \( h_2 \), des fokussierten Sachverhalts weder in einer parallelen noch einer orthogonalen Lagebeziehung (vgl. Abbildung 4.42). Das \( PRS \) ist wie bisher durch die Hauptbewegungsrichtung entlang des Pfades gegeben. Das \( SRS \) leitet sich vom Rotationswinkel \( \beta \) zwischen der Frontalachse des \( PRS \) und dem Schwerpunkt des räumlichen Sachverhalts (Landmarken \( A \) und \( C \)) ab (vgl. Abbildung 4.42). Potentiell sind drei virtuelle Referenzsysteme \( VRS_1 \), \( VRS_2 \) und \( VRS_3 \) etablierbar. \( VRS_2 \) und \( VRS_3 \) werden

![Abbildung 4.40: Orientierung eines virtuellen Referenzsystems](image-url)
Abbildung 4.41: Virtuelle Referenzsysteme \((VRS_1(t), VRS_2(t))\) zum Zeitpunkt \(t\) bezüglich der Hauptachsen \(h_1\) und \(h_2\).

Abbildung 4.42: Virtuelle Referenzsysteme bei geneigten Konfigurationen

bezüglich der beiden Hauptachsen \(h_1\) und \(h_2\) des Referenzobjektes \(A\) an den Lokationen \(l_2\) und \(l_3\) definiert. Da zwischen dem Agenten und dem räumlichen Sachverhalt eine große Distanz besteht, sind die beiden Winkel \(\delta\) und \(\gamma\) groß. In diesem Fall bietet sich aus kognitiv-ökonomischen Gründen eine alternative Verankerungsmöglichkeit durch Rückgriff auf das primäre Referenzsystem an. \(VRS_I\) wird in einer Lokation \(l_I\) verankert, die sich auf einer durch den Schwerpunkt des Referenzobjektes gehenden Linie parallel zur Bewegungsrichtung befindet. Es lassen sich somit für die Verankerung eines geeigneten virtuellen Referenzsystems im allgemeinen zwei Fälle unterscheiden.

1. Verankerung bezüglich der Hauptachsen des Referenzsystems
2. Verankerung bezüglich der Hauptachse des primären Referenzsystems

Nicht in jedem Fall kann vom Referenzobjekt in dieser Weise abstrahiert werden, daß die Information über den Schwerpunkt und die Hauptachsen ausreichen, was im folgenden behandelt wird.

4.2.4.1.6 Berticksichtigung intrinsicher Eigenschaften bei der Wahl eines virtuellen Referenzsystems In vielen Fällen besitzt das Referenzobjekt ausgezeichnete intrinsische
4.2. FUNKtionale Beschreibung

Eigenschaften, die die Auswahl eines geeigneten virtuellen Referenzsystems beeinflussen (vgl. Kapitel 2.3.1). Nimmt man beispielsweise an, daß das Referenzobjekt $A$ in Abbildung 4.42 eine Kirche ist, dann ist die Front der Kirche durch ihr Portal gegeben. Die Objektklasse der Kirchen besitzt die intrinsische Eigenschaft, ein virtuelles Referenzsystem zu prädeterminieren. Wie beispielsweise nachgewiesen wurde können, wird in der deutschen Sprache bezüglich einer Kirche ein virtuelles Referenzsystem entsprechend $\text{VRS}_5$ bevorzugt (vgl. [Car93]). Ein Betrachter nimmt somit eine Perspektive ein, als ob er von außen auf das Portal der Kirche schauen würde. Solche intrinsischen Eigenschaften sind funktionale, soziale, kulturelle oder ähnliche Merkmale, die einem einzelnen oder einer Klasse von Objekten zugeordnet werden.


![Abbildung 4.43: Verankerung eines intrinsisch determinierten, virtuellen Referenzsystems ($\text{VRS}_{int}$) in einem Referenzobjekt (Automobil) aufgrund prototypischen Wissens über den Umgang mit diesem.](image)

Intrinsisch determinierte virtuelle Referenzsysteme sind weitgehend von Rotationswinkeln unabhängig. Aufgrund der Untersuchungen von Herrmann et al. ist zu vermuten, daß dies daher rührt, daß der kommunikative Akt einer Raumbeschreibung durch ein solches prototypisches Wissen und durch Konventionen erheblich vereinfacht wird. Eine solche Erklärung läßt sich auch auf Adressaten einer Beschreibung ausweiten. Wie in mehreren Untersuchungen gezeigt worden ist, werden Beschreibungen in kommunikativen Kontexten präferiert, in denen der Sprecher die Position des Rezipienten einnimmt (vgl. [HBN87], [Sch95b], [Som69]).

\textsuperscript{46}Interessanterweise etablieren englischsprachliche Personen das virtuelle Referenzsystem so, als ob sie aus dem Portal hinausschauen würden. Dadurch bedingt ist eine Spiegelung dessen, was als „links“ von und „rechts“ von der Kirche bezeichnet wird.

\textsuperscript{46}Interessant sind Fälle, bei denen man nicht allgemein für eine Objektklasse bestimmen kann, ob das $\text{VRS}$ extern oder intern zum Referenzobjekt verankert wird. Als Beispiel sei die Klasse der Gebäude genannt. Spricht man über das Elternhaus, so ist das $\text{VRS}$ typischerweise intern verankert. Für den Deutschen Reichstag hingegen wird typischerweise eine externe Verankerung verwendet.
Zusammenfassend lassen sich mit primären, sekundären und virtuellen drei verschiedene Klassen von Referenzsystemen identifizieren, die in dynamischen Situationen grundlegend für die Verankerung raumbezogenen Wissens sind. Primäre und sekundäre Referenzsysteme sind egozentrisch, wohingegen virtuelle Referenzsysteme allozentrisch sind. Zur Auswahl eines geeigneten virtuellen Referenzsystems wird eine extrinsische und eine intrinsische Strategie verwendet. Die extrinsische bezieht sich einzig auf geometrische Eigenschaften des Referenzobjektes, wie Hauptachsen und Schwerpunkte. Die intrinsische Strategie basiert auf intrinsischen Eigenschaften des betrachteten Objektes, die aus verschiedenen Quellen stammen können und deren Formalisierung i.a. komplex ist.

4.2.4.2 Evaluation räumlicher Relationen


Bei der Evaluation von räumlichen Relationen lassen sich mehrere Ebenen unterscheiden. In den meisten Arbeiten wird eine Unterteilung in zwei Ebenen angenommen, die das Konzept eines Objektes in geometrisch- strukturelle und semantische Eigenschaften zerlegen (vgl. Kapitel 2.3.2). Die Evaluation räumlicher Relationen, die auch hier verwendete, basiert nahezu ausschließlich auf einer geometrisch- strukturellen Repräsentationsebene. Semantische Eigenschaften von Objekten spielen nur eine eingeschränkte Rolle (vgl. [Her86], [LCS90], [Gap94]).

In diesem Kapitel werden die drei verschiedenen Relationenarten (topologisch, deiktisch und intrinsisch) nacheinander soweit diskutiert, wie sie für die Verwendung in dieser Arbeit notwendig sind.

4.2.4.2.1 Prinzip zur Bestimmung von Anwendbarkeitsgraden über Lokationen

Räumliche Relationen sind im sprachlichen Gebrauch nicht durch zweiwertige Logikprädikate definierbar. Vielmehr sind räumliche Relationen umscharfe Konzepte, deren Verwendung gradu ell Unterschiede aufweisen (z.B. [Her86], [Han80]). Auf sprachlicher Ebene sind linguistische Hecken Markierungen solcher Vagheiten (vgl. [Lak72]). Zur Formalisierung werden die Gebiete um das Referenzobjekt ro in diskrete räumliche Lokationen47 unterteilt (z.B. [Her96], [Gap94], [OMT94], [Hab87]). Ein Beispiel eines solche gebiets konstituierenden Ansatzes ist das in VITRA entwickelte räumliche Relationenmodell (vgl. [ABHR87], [RS88], [AHR89], [Gap94]). In diesem Modell wird einer jeden Relation über eine Anwendbarkeitsfunktion eine Möglichkeit zugeordnet, mit der diese Relation auf sprachlicher Ebene direkt als sprachliche Präposition verwendet wird (vgl. [SBS87]). Jeder Lokation ist dadurch bezüglich einer Anwendbarkeitsfunktion eine Graduierung auf dem Intervall [0...1] zugeordnet, die angibt, inwieweit eine assoziierte räumliche Präposition den betrachteten räumlichen Sachverhalt zwischen dem Referenzobjekt und dem Lokationsobjekt to beschreibt. Als Eingabe gehen in diese Funktionen die Größe des Referenzobjektes sowie die Distanz und die Winkelabwärzung des Lokationsobjektes zum Referenzobjekt mit ein (vgl. [Gap94]). Basierend auf der These von Landau und Jackendorf (vgl. [LJ93]) wird in

47Räumliche Lokationen sind dabei entweder zwei- oder null-dimensional.
diesem Modell davon ausgegangen, daß nur stark abstrahierte geometrische Eigenschaften des Referenzobjektes und des Lokationsobjektes in die Evaluation räumlicher Relationen eingehen. Das Referenzobjekt wird dabei durch einen minimal-umschreibenden Quader und lo durch seinen Schwerpunkt repräsentiert.48 Die Größe des Referenzobjektes geht durch eine Normierung des Koordinatensystems ein (vgl. [Gap94]). Dabei wird angenommen, daß der Raum homogen ist, d.h., daß es keine ausgezeichneten Lokationen im Raum gibt und daß der Raum verzerrungsfrei bezüglich einer linearen Metrik ist.49 Unter dieser Voraussetzung kann eine Situation bezüglich der Größe des Referenzobjektes normiert werden.50 Einem solchen Ansatz liegt eine vollständige geometrische Modellierung einer Situation zugrunde, von der sich Distanzen und Winkelabweichungen direkt ableiten lassen. Inwieweit eine solche vollständige Repräsentation kognitive Plausibilität besitzt, ist eine offene Frage. In dem Modell werden ausschließlich statistische Situationen betrachtet (z.B. [Gap94]). Eine Bezugsnahme auf die Lokation des Sprechers wird dabei nur implizit verwendet, weswegen eine räumliche Relation R allgemein als binär zwischen dem Lokalisations- und Referenzobjekt definiert ist (R(lo, ro), z.B. [MJJ76]). Bewegt sich der Sprecher jedoch, ändern sich gleichzeitig über die Lokation des Agenten die egozentrischen und virtuellen Referenzsysteme. Über Referenzsysteme bestimmt sich, welche Art der Beschreibung (deiktisch, intrinsisch oder extrinsisch) verwendet wird. Dadurch geht die Bewegung des Agenten direkt in die Berechnung räumlicher Relationen mit ein, weswegen im weiteren räumliche Relationen als dreistellig verstanden werden: R(ego, lo, ro). Nacheinander wird im weiteren diskutiert, in welcher Weise topologische, deiktische und intrinsische räumliche Relationen in dieser Arbeit verwendet werden. Für weitergehende Ausführungen zum Konzept der räumlichen Relationen sei auf Kapitel 2.3.2 und entsprechende Arbeiten verwiesen (vor allem [MJJ76], [Her86], [Pri93], [Her96], [Gap94]).

4.2.4.4.2 Topologische räumliche Relationen Topologische räumliche Relationen sind unabhängig von Referenzsystemen. In der hier betrachteten Domäne der Wegbeschreibungen sind sie nur insoweit wichtig, daß sie Kontaktrrelationen zwischen Pfadelementen repräsentieren. Um sicherzustellen, daß ein Agent direkt von einem Pfadelement zum nächsten gelangen kann, muß mindestens eine Kontaktrrelation zwischen beiden bestehen. Legt man eine Punktmenge-repräsentation der Objekte zugrunde, so liegt zwischen zwei Objekten A und B dann eine Kontaktrrelation vor, wenn wenigstens ein Punkt sowohl in A, als auch in B ist. In diesem Falle gilt die Relation \( \#meet\#(\text{loc}(A), \text{loc}(B)) \)).51 Eine solche Definition einer Kontaktrrelation beruht auf einer Formalisierung von topologischen Relationen im zwei- und dreidimensionalen Raum (vgl. Kapitel 2.3.2, insb. [Ege91], [EAT92], [Rup96]). Danach lassen sich sechs topologische Relationen unterscheiden (vgl. Abbildung ??).

topo-rel]
Die Evaluation von topologischen Relationen zwischen zwei zweidimensionalen Objekten \( \text{obj}_j \)

---

48 Dies gilt aber nur unter der Annahme, daß die Objekte konvex sind.
50 Über diese Modellierung wird das Phänomen repräsentiert, daß sich bezüglich größerer Referenzobjekte räumliche Relationen über einen größeren räumlichen Bereich als anwendbar erweisen als bei kleineren.
51 Anstelle von \( \text{loc}(A) \) wird im weiteren abkürzend auf die explizite Nennung der Funktion \( \text{loc} \) verzichtet.
und obj2 erfolgt durch die Funktion *Evaluate-Topological-Rel* (vgl. Definition 23).52

\[(23)\quad \text{Evaluate-Topological-Rel}(\text{loc}(\text{obj}_1), \text{loc}(\text{obj}_2));\]

### 4.2.4.2.3 Deiktische Relationen

Die Evaluation einer deiktischen Relation basiert entweder auf einem primären oder einem sekundären Referenzsystem, welches in der Lokation des Agenten (ego) verankert ist. Entsprechend der Hauptachsen eines kartesischen Referenzsystems werden sechs deiktische Relationen unterschieden (vgl. [Her86], [Log95], [HT95]):

1. #left-of#(ego, lo, ego): lo befindet sich links vom Agenten
2. #right-of#(ego, lo, ego): lo befindet sich rechts vom Agenten
3. #behind#(ego, lo, ego): lo befindet sich hinter dem Agenten
4. #in-front-of#(ego, lo, ego): lo befindet sich vor dem Agenten
5. #above#(ego, lo, ego): lo befindet sich über dem Agenten
6. #under#(ego, lo, ego): lo befindet sich unter dem Agenten

Sind bezüglich eines neuen Objekts deiktische räumliche Relationen zu bestimmen, so ist das Referenzobjekt der Agent. Deiktische räumliche Relationen repräsentieren die Lage eines Objektes relativ zur Hauptausrichtung des Agenten, weswegen sie bezüglich eines primären oder sekundären Referenzsystems evaluiert werden.53 Die Eingabe der entsprechenden Funktion zur Bestimmung der bestanwendbaren deiktischen Relation ist die Funktion *Evaluate-Deictic-Rel*.

---

52 Eine dedizierte Analyse der formalen Berechnung topologischer Relationen liegt außerhalb des Kontext dieser Arbeit. Verschiedene Berechnungsmodelle topologischer Relationen sind in Kapitel 2.3.2 referenziert.

53 Zur Evaluation von räumlichen Relationen basierend auf geometrischen Objektrepräsentationen (vgl. [Her96]).
Abbildung 4.45: Entsprechend der Funktion \( \text{right-of}(\text{ego}, \text{lo}, \text{ego}) \) konstituiertes Gebiet der Anwendbarkeitsgrade

Einer Illustration des Zusammenhangs zwischen Objekten, Referenzsystemen und Anwendbarkeitsgraden im deiktischen Falle, dient die Abbildung 4.45. Das Gebiet der Anwendbarkeitsgrade ist durch den schattierten Bereich um die Lokation des Betrachters indiziert. Dunkle Gebiete indizieren hohe, helle geringere Anwendbarkeitsgrade. Im Falle der deiktischen Evaluation ist das Anwendbarkeitsgebiet im PRS bzw. im SRS verankert. Im vorliegenden Fall besteht eine hohe Anwendbarkeit der Relation \( \text{right-of} \) für das Objekt \( \text{lo} \).

### 4.2.4.2.4 Intrinsische räumliche Relationen

Die Evaluation intrinsicher räumlicher Relationen setzt die Etablierung eines virtuellen Referenzsystems voraus. Die Menge der hier betrachteten intrinsischen Relationen stimmt mit den deiktischen Relationen überein, mit dem Unterschied, daß das Referenzsystem sich an einer externen Lokation befindet:\(^{54}\)

1. \( \text{left-of}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich links vom Referenzobjekt
2. \( \text{right-of}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich rechts vom Referenzobjekt
3. \( \text{behind}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich hinter dem Referenzobjekt
4. \( \text{in-front-of}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich vor dem Referenzobjekt
5. \( \text{above}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich über dem Referenzobjekt
6. \( \text{under}(\text{ego}, \text{lo}, \text{refo}) \): \( \text{lo} \) befindet sich unter dem Referenzobjekt

\(^{54}\)Der hier verwendete Ansatz erlaubt es, nur solche intrinsische Relationen zu berechnen, die bezüglich einem im Referenzobjekt verankerten virtuellen Referenzsystem definiert sind (vgl. [Gap94]).

\[(25) \quad \text{Evaluate-Intrinsic-Rel}s(\text{ego}, \text{lo}, \text{refo})\]

Abbildung 4.46: Evaluation höchstanwendbarer intrinsischer Relationen: a) bei antiparallelem VRS, b) bei orthogonal-orientiertem VRS

Durch die einzelsprachabhängige Festlegung des virtuellen Referenzsystems, ergibt die Evaluation intrinsischer Relationen unterschiedliche Relationen mit höchstem Anwendbarkeitsgrad (vgl. Abbildung 4.46). Im ersten Fall (vgl. Abbildung 4.46a)) ist das virtuelle Referenzsystem antiparallel zum primären bzw. sekundären Referenzsystem des Betrachters. Die höchstanwendbare intrinsische Relation ist dabei \#in-front-of#. Im zweiten Fall (Abbildung 4.46b) ist das VRS nach rechts orientiert, so daß in dieser Situation die Relation \#right-of# den höchsten Anwendbarkeitsgrad besitzt.

Zusammenfassend dienen die hier betrachteten räumlichen Relationen der Integration von einzelnen Objekten zu einer propositionalen Repräsentation eines räumlichen Sachverhalts. Topologische Relationen werden dazu verwendet, um physikalische Kontaktbeziehungen zu repräsentieren, wobei im weiteren alleinig die Relation \#meet# verwendet wird. Deiktische und intrinsische Relationen basieren im Gegensatz zu topologischen Relationen auf der Verwendung eines Referenzsystems. Die räumlichen Beziehungen zwischen einem einzelnen Objekt und dem

---

55In diesem Beispiel wird zur besseren Illustration davon ausgegangen, daß das virtuelle Referenzsystem durch intrinsische Eigenschaften des Referenzobjektes in eine antiparallele bzw. orthogonale Orientierung relativ zum PRS/SRS rotiert wird. Eine realweltliche Instantiierung dieses Beispiels ist gegeben, wenn das Referenzobjekt eine Kirche mit Portal an der Seite zum Lokalisationsobjekt ist (vgl. Abbildung 4.46a), bzw. wenn sich das Portal an der rechten Seite befindet (vgl. Abbildung 4.46b).

Betrachtet werden auf der Basis eines egozentrischen Referenzsystems durch deiktische Relationen repräsentiert. Hingegen werden intrinsische Relationen bezüglich eines allozentrischen, virtuellen Referenzsystems, verwendet, um räumliche Lagebeziehungen zwischen zwei Objekten zu repräsentieren.


4.2.4.3 Der räumliche Kurzzeitspeicher

In den vorangegangenen Abschnitten ist vorgestellt worden, wie Pfadelemente gesucht (vgl. Kapitel 4.2.2), visuell saliente Objekte ausgewählt (vgl. Kapitel 4.2.3), sowie Referenzsysteme (vgl. Kapitel 4.2.4.1) und räumliche Relationen (vgl. Kapitel 4.2.4.2) etabliert werden. Im weiteren wird diskutiert, wie alle diese Einzelkonzepte integriert werden, um eine räumliche Vorstellung einer Situation, eine sogenannte räumliche Konfigurationsbeschreibung, zu etablieren. Bedingt durch die Dynamik einer Situation, wird eine solche räumliche Konfigurationsbeschreibung inkrementell evaluiert und angepaßt. Das dadurch etablierte raumbezogene Wissen über eine Situation wird im sogenannten räumlichen Kurzzeitspeicher verwaltet.


4.2.4.3.1 Räumliche Konfigurationsbeschreibungen Um in dynamischen Situationen Raumbeschreibungen generieren zu können, muß der Agent über lokal-stabiles, raumbezogenes Wissen einer Situation verfügen. Entsprechend dem Ansatz, der zur Auswahl salienter Objekte vorgestellt worden ist (vgl. Kapitel 4.2.3), wird raumbezogenes Wissen bezüglich diskreten

Abbildung 4.47: Zweistufige Generierung raumbezogenen Wissens in dynamischen Situationen

Eine räumliche Konfigurationsbeschreibung (RKB) repräsentiert auf qualitative Weise visuell-räumliche Sachverhalte einer Situation. Gebietskonstituierende, räumliche Relationen, so wie sie in Kapitel 4.2.4.2 vorgestellt worden sind, sind weitgehend invariant bezüglich geringer räumlicher Lageänderungen, wodurch sich eine Stabilität einer RKB ergibt. Bewegt sich der Agent translatorisch, so selektiert er neue Entitäten aus der sich ändernden Situation. Entitäten, die nicht mehr sichtbar sind, werden zusammen mit den räumlichen Relationen aus der RKB entfernt (vgl. Kapitel 3).

Sobald der Agent eine neue Situation erreicht, wird die RKB initialisiert. Dies ist eine Einschränkung der Allgemeingültigkeit bezüglich der Verarbeitung räumlichen Wissens. Hingegen ist dies in der Domäne von inkrementalen Wegbeschreibungen legitim, da in keiner der Beschreibungen aus den Untersuchungen Beziehe zu räumlichen Wissen aus vorangegangen Pfadabschnitten gemacht worden sind.

4.2.4.3.1.1 Minimale räumliche Konfigurationsbeschreibungen Jeder Verzweigungspunkt beendet den aktuell verfolgten Pfadabschnitt und veranlasst den Agenten, sich und dadurch auch das primäre Referenzsystem in eine bestimmte Richtung zu drehen, um dann auf diesem Pfadabschnitt die Navigation fortzusetzen. Zwischen zwei Pfadabschnitten befindet sich grundsätzlich ein Verzweigungspunkt (vgl. Kapitel 4.2.2). Eine physische Grundbedingung an Pfadelemente ist, daß der Übergang zwischen einem Pfadabschnitt und einem Verzweigungspunkt stetig ist und eine Kontaktrelation erfüllt ist, die durch die topologische Relation "meet" repräsentiert wird. Die Integration eines Pfadelementes in eine räumliche Konfi-

---

57 Eine natürliche Fortsetzung dieser Arbeit ist es, das Wissen einer räumlichen Konfigurationsbeschreibung durch geeignete Funktionen an eine Langzeitspeicherung in Form einer kognitiven Karte weiterzuleiten (vgl. Kapitel 2.3.7, 26).
58 Die Menge der Pfadelemente ist die Gesamtheit aus Pfadabschnitten und Verzweigungspunkten (vgl. Kapitel 4.2.2).
59 Der Aspekt der Stetigkeit wird im weiteren Verlauf nicht weiter betrachtet. Ein geeignetes Konzept der qualitativen Stetigkeit ist eng mit dem Energieaufwand verknüpft, der den Agent benötigt, um den Übergang erfolgreich zu überwinden. Sprünge bei Übergängen, wie bei Treppen, müssen bis zu einem gewissen Grade als qualitativ stetig konzeptualisiert werden.
Etablierung deikt. Relationen

Etablierung topolog. Relationen

Etablierung intrins. Relationen

Abbildung 4.48: Stufen zur Integration einer Pfadelemente

In Abbildung 4.49 ist eine Situation angegeben, in der der Agent nach links in den Pfadabschnitt S2 abbiegen möchte. Hierzu werden deiktische Relationen zwischen den relevanten Pfadelementen S1, C, und S2 etabliert. Da sich der Agent zu diesem Zeitpunkt bereits auf S1 bewegt, gilt, daß sich S1 „unter“ ego befindet (vgl. Relationen in 26).

Abbildung 4.49: a) Situation mit einer Intention des Linksabbiegens und b) deiktische Relationen bezüglich der Pfadelemente

\[(26) \quad \#_{\text{under}}(\text{ego}, \text{S1}, \text{ego})\]

Zwischen dem Pfadabschnitt S2 und dem Agenten besitzen zwei deiktische Relationen einen hohen Anwendbarkeitsgrad, die als Komposition ergeben, daß sich S2 „links vor“ dem Agenten
befindet (vgl. Relationen in 27).

\[(27) \quad \#\text{left-of}#(\text{ego, S2, ego}) \& \#\text{in-front-of}#(\text{ego, S2, ego})\]

Zum Verzweigungspunkt \( C \) läßt sich wiederum nur eine deiktische Relation etablieren, die repräsentiert, daß \( C \) sich „vor“ dem Agenten befindet (vgl. Relationen in 28).

\[(28) \quad \#\text{in-front-of}#(\text{ego, C, ego})\]

Zur Evaluierung physikalischer Begegnheiten wird zwischen dem Agenten und dem Pfadabschnitt \( S1 \) evaluiert, ob eine Kontak트relation besteht (vgl. Abbildung 4.50a) und die Relationen in 29).

\[(29) \quad \#\text{meet}#(\text{ego, S1})\]

Weiterhin bestehen Kontak트relationen zwischen den Pfadelementen (vgl. Relationen in 30).

\[(30) \quad \#\text{meet}#(\text{S2, C}), \#\text{meet}#(\text{C, S1})\]

Abbildung 4.50: a) Topologische Relationen, b) intrinsische Relationen zwischen Pfadelementen, c) minimale räumliche Konfigurationsbeschreibung

Das Wissen, was alleine durch deiktische und topologische Relationen repräsentiert wird, ist als Grundlage für sprachliche Beschreibungen nicht ausreichend, da u. a. die Transitivitysei-
4.2. **FUNKTIONALE BESCHREIBUNG**


\[(31) \text{ #left-of#(ego, S2, S1) \& #behind#(ego, S2, S1)}\]

Ferner gilt, daß aus Sicht des Agenten die Kreuzung C „hinter“ S1 und das S2 „links von“ C liegt (vgl. Relationen in 32).

\[(32) \text{ #behind#(ego, C, S1) \& #left-of#(ego, S2, C)}\]

Insgesamt besteht eine minimale räumliche Konfigurationsbeschreibung aus der Gesamtheit der räumlichen Relationen mit den höchsten Anwendbarkeitsgraden, so wie es in Abbildung 4.50c) indiziert ist.

### 4.2.4.3.1.2 Pfadgeleitete Intentionsmuster

Aus den intrinsischen Relationen zwischen Pfadelementen lassen sich pfadgeleitete Intentionen des Agenten bezüglich einer Situation S ableiten. Durch die qualitative räumliche Repräsentation der Pfadelemente zueinander lassen sich **Intentionsmuster (IM)** angeben, nach denen entschieden werden kann, in welche Richtung der Pfad weiter verfolgt wird. Basierend auf den intrinsischen Relationen zwischen Pfadelementen konstituieren sich sechs Intentionsmuster, denen **pfadgeleitete Intentionen (PI)** zugeordnet sind (vgl. Abbildung 4.51). Entsprechend den kartesischen Dimensionen der Referenzsysteme werden folgende sechs\(^{61}\) pfadgeleitete Intentionen unterschieden:

1. **nach-rechts-bewegen**: der nachfolgende Pfadabschnitt S2 liegt qualitativ „rechts“ vom Verzweigungspunkt C und dem aktuellen Pfadabschnitt S1 ($\text{turn-right}(S1, C, S2)$).\(^{62}\)
2. **nach-links-bewegen**: S2 liegt qualitativ „links“ von C und S1 ($\text{turn-left}(S1, C, S2)$).
3. **geradeaus-bewegen**: S2 liegt qualitativ „hinter“ C und S1 ($\text{go-straight}(S1, C, S2)$).
4. **nach-oben-bewegen**: S2 liegt qualitativ „über“ C und S1 ($\text{go-up}(S1, C, S2)$).
5. **nach-unten-bewegen**: S2 liegt qualitativ „unter“ C und S1 ($\text{go-down}(S1, C, S2)$).
6. **horizontal-bewegen**: S2 liegt qualitativ „eben“ bezüglich C und S1 ($\text{go-plane}(S1, C, S2)$).

Da intrinsische Relationen zwischen Pfadelementen relativ zum primären Referenzsystem etabliert sind, folgt eine ebensolche Verankerung pfadgeleiteter Intentionen. Die Bestimmung einer PI beruht auf einer schwellwertbasierten Auswertung des assoziierten Intentionsmusters (vgl.

---

\(^{60}\) Liegt ein Pfadelement A links sowie die Elemente B und C vor dem Agenten, so kann nicht eindeutig abgeleitet werden, wie A bezüglich B liegt.

\(^{61}\) Die Möglichkeit einer Rückwärtsbewegung wird nicht betrachtet.

\(^{62}\) Pfadgeleitete Intentionsmuster sind durch $\text{prä-}$ und postfigiert.
Abbildung 4.51: Intensionsmusterbasierte Ableitung pfadgeleiteter Intentionen: a) linksabbiegen, b) rechtsabbiegen, c) geradeausfahren, d) hochfahren, e) hinunterfahren und f) ebenfahren

Definition 33). Über Schwellwerte wird festgelegt, wie groß die Abweichung von einer prototypischen räumlichen Anordnung von S1, C und S2 sein darf. Die Berechnung des Anwendbarkeitsgrades \( AG_{int} \) einer PI \( (AG_{int}(IM, S)) \) in einer Situation S erfolgt per Durchschnittsbildung über die Anwendbarkeitsgrade der intrinsischen Relationen \( (AG_{rel}) \) der Intensionsmuster IM (vgl. Definition 33).

\[
(33) \quad AG_{int}(IM, S) = \frac{1}{n} \sum_{i=1}^{n} AG_{rel}(rrfi(IM, S))
\]

Abbildung 4.52: a) Prototypische Situation eines $\text{turn-left}$ Intensionsmusters, b) und c) nicht-prototypische Situationen eines $\text{turn-left}$ Intensionsmusters

Die bisher vorgestellten Intentionen lassen sich entweder einer Klasse \( PI_{horizontal} \), die auf die horizontale (vgl. Abbildung 4.51a)-c)) oder einer Klasse \( PI_{vertical} \), die auf die vertikale Ebene
(vgl. Abbildung 4.51(d-f) und Gleichung 34) bezogen ist, zuordnen.

\[
\begin{align*}
\text{PI}_{\text{horizontal}} &= \{\text{turn-left\$}, \text{turn-right\$}, \text{go-straight\$}\} \\
\text{PI}_{\text{vertikal}} &= \{\text{go-up\$}, \text{go-down\$}, \text{go-plane\$}\}
\end{align*}
\]


<table>
<thead>
<tr>
<th>Konjungierte PI</th>
<th>Konjunktion linear unabhängiger PIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. turn-left$</td>
<td>turn-left$ ∧ go-plane$</td>
</tr>
<tr>
<td>2. turn-left-up$</td>
<td>turn-left$ ∧ go-up$</td>
</tr>
<tr>
<td>3. turn-left-down$</td>
<td>turn-left$ ∧ go-down$</td>
</tr>
<tr>
<td>4. turn-right$</td>
<td>turn-right$ ∧ go-plane$</td>
</tr>
<tr>
<td>5. turn-right-up$</td>
<td>turn-right$ ∧ go-up$</td>
</tr>
<tr>
<td>6. turn-right-down$</td>
<td>turn-right$ ∧ go-down$</td>
</tr>
<tr>
<td>7. go-straight$</td>
<td>go-straight$ ∧ go-plane$</td>
</tr>
<tr>
<td>8. go-up$</td>
<td>go-straight$ ∧ go-up$</td>
</tr>
<tr>
<td>9. go-down$</td>
<td>go-straight$ ∧ go-down$</td>
</tr>
</tbody>
</table>

Tabelle 4.2: Konjunktionen von linear unabhängigen pfadgeleiteten Intentionen zu konjungierten pfadgeleiteten Intentionen

Durch Konjunktion von pfadgeleiteten Intentionen aus PI_{horizontal} und PI_{vertikal} lassen sich neu konjungierte pfadgeleitete Intentionen ableiten (vgl. Tabelle 4.2). Wie aus Tabelle 4.2 ersichtlich ist, besitzen die Klassen PI_{horizontal} mit go-straight\$ und PI_{vertikal} mit go-plane\$ neutrale Elemente. Werden diese beiden PIs miteinander konjungiert (Zeik 7 in Tabelle 4.2), so dominiert die Intention aus PI_{horizontal} Per Durchschnittsbildung des Anwendbarkeitsgrades von PI (AG_{int}(IM_1, S)) und PI (AG_{int}(IM_2, S)), ergibt sich der Anwendbarkeitsgrad der konjungierten pfadgeleiteten Intention (KPI) AG_{kpi}(IM_1, IM_2, S) (vgl. Gleichung 35).

\[
AG_{kpi}(IM_1, IM_2, S) = \frac{1}{2}(AG_{int}(IM_1, S) + AG_{int}(IM_2, S))
\]


### 4.2.4.3.1.3 Erweiterte räumliche Konfigurationsbeschreibungen

In der zweiten Phase des Aufbaus raumbezogenen Wissens (vgl. Abbildung 4.47) werden Landmarken integriert, die aufgrund ausreichender visueller Salienz ausgewählt worden sind (vgl. Kapitel 4.2.3). Landmarken dienen als räumliche Referenzpunkte in einer Situation und sind von pfadgeleiteten
Intentionen abhängig.


Abbildung 4.53: Konstruktion eines sekundären Referenzsystems in Abhängigkeit von einer pfadgeleiteten Intention des Linksabbiegens


\[(36) \quad \#\text{left-of} \#(\text{ego, } A, \text{ ego}) \& \#\text{in-front-of} \#(\text{ego, } A, \text{ ego})\]

\(^{\text{63}}\)Solche \(PIs\) enthalten in ihrer Definition \$\text{turn-left}\$ oder \$\text{turn-right}\$. 
4.2. FUNKTIONALE BESCHREIBUNG

Abbildung 4.54: Stufen zur Integration einer Landmarke

Funktion 10: Integrate-Landmarks

\[ \text{Integrate-Landmarks}(S, PRS, focus, lm) \]
\[ \text{Evaluate-Deictic-Rel}(SRS, lm); \]
\[ \text{vs} := \text{Select-Virtual-Frame-of-Reference}(srs, lm); \]
\[ \text{Determine-Spatially-Near-Objects}(lm); \]
\[ \text{for all} \ o \ \text{in} \ \text{RAGM}(lm) \ \text{do} \]
\[ \text{Evaluate-Intrinsic-Rel}(vs, lm, o); \]
\[ \text{Det-Appl-of-LM-in-RKB}(lm) \]


Die bezüglich Abbildung 4.53 evaluierten räumlichen Relationen sind in Abbildung 4.55 eingezeichnet.\footnote{Es werden der Klarheit wegen für Landmarken extrinsisch determinierte, virtuelle Referenzsysteme angenommen.} Danach befindet sich der Pfadabschnitt \textit{S1} rechts von \textit{A}, der Verzweigungspunkt \textit{C} rechts hinter \textit{A} und der Pfadabschnitt \textit{S2} links hinter \textit{A} (vgl. die intrinsischen Relationen 37).

\begin{equation}
(37) \quad \#\text{right-of}(\text{ego, A, S1})
\end{equation}
\begin{equation}
\#\text{right-of}(\text{ego, A, C}) \& \#\text{behind}(\text{ego, A, C})
\end{equation}

Abbildung 4.55: a) Intrinsische Relationen zwischen Pfadelementen und der Landmarke A und b) erweiterte räumliche Konfigurationsbeschreibung


Abbildung 4.56: Mögliche Lokation eines potentiellen Referenzobjektes bei den pfadbezogenen Intentionen a) $\texttt{turn-left}$ und b) $\texttt{turn-right}$. 
4.2. FUNKTIONALE BESCHREIBUNG

Die Berechnung der Anwendbarkeit einer Landmarke erfolgt musterbasiert. Aufbauend auf den möglichen Lokationen potentieller Referenzobjekte, lassen sich für jede Abbiegeintention die obengenannten zwei Fälle unterscheiden, die verschiedene räumliche Anwendbarkeitsgrad muster (RAGM) besitzen (vgl. Tabelle 4.3). Landmarken, die sich vor S2 befinden, sind durch $L_{\text{front}}$ und solche hinter S2 durch $L_{\text{back}}$ indiziert.

<table>
<thead>
<tr>
<th>RAGM($$\text{turn-left}$)</th>
<th>S1</th>
<th>C</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{\text{front}}$</td>
<td>#right-of#</td>
<td>#right-of#</td>
<td>#behind#</td>
</tr>
<tr>
<td>$L_{\text{back}}$</td>
<td>#right-of#</td>
<td>#in-front-of#</td>
<td>#in-front-of#</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RAGM($$\text{turn-right}$)</th>
<th>S1</th>
<th>C</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{\text{front}}$</td>
<td>#left-of#</td>
<td>#left-of#</td>
<td>#behind#</td>
</tr>
<tr>
<td>$L_{\text{back}}$</td>
<td>#left-of#</td>
<td>#in-front-of#</td>
<td>#in-front-of#</td>
</tr>
</tbody>
</table>

Tabelle 4.3: Musterbasierte Auswahl von potentiellen Referenzobjekten

![Diagram](image)

Abbildung 4.57: Zwei potentielle Referenzobjekte A und B

Die Evaluation der Anwendbarkeit einer Landmarke erfolgt über die räumlichen Anwendbarkeitsgrade der einzelnen räumlichen Relationen. Der räumliche Anwendbarkeitsgrad RAG(lm) einer Landmarke ergibt sich aus dem arithmetischen Mittel der Quadrate räumlicher Anwendbarkeitsgrade der räumlichen Relationen, die durch das betreffende räumliche Anwendbarkeitsgrad-
muster vorgegeben sind (vgl. Gleichung 38). Ein geeignetes $n$ ergibt sich aus der anwendbaren Eintrag aus den Tabellen 4.3. Befindet sich die Landmarke vor dem nächsten Pfadabschnitt, so ist $n$ gleich vier, liegt sie dahinter, so ist $n$ gleich 3 (vgl. Tabelle 4.3).

\[
(38) \quad \text{RAG}(\text{lm}) = \frac{1}{n} \sum_{i=1}^{n} \text{rel}(R_i(\text{ego, lm, obj}i))^2
\]

In Abbildung 4.57 erfüllen beide Landmarken $A$ und $B$ die Bedingung in der geeigneten Halbebene zu liegen. Beide Landmarken liegen vor dem nächsten Pfadabschnitt $S2$, so daß das erste räumliche Anwendbarkeitsgradmuster von $\text{RAG}(\text{turn-left})$ anwendbar ist. Wie sich aus den folgenden Berechnungen ergibt, besitzt die Landmarke $A$ einen höheren räumlichen Anwendbarkeitsgrad $\text{RAG}$ als $B$.

\[
(39) \quad \text{RAG}(A) = (\#\text{right-of}(\text{ego, A, S1})^2 + \#\text{right} - \#\text{of}(\text{ego, A, C})^2 + \#\text{behind}(\text{ego, A, C})^2 + \#\text{behind}(\text{ego, A, S2})^2) / 4
\]

\[
= (1^2 + 0, \overline{7}^2 + 0, \overline{7}^2 + 0, 0) / 4 = 0, 6975
\]

\[
(40) \quad \text{RAG}(B) = (\#\text{right-of}(\text{ego, B, S1})^2 + \#\text{right} - \#\text{of}(\text{ego, B, C})^2 + \#\text{behind}(\text{ego, B, C})^2 + \#\text{behind}(\text{ego, B, S2})^2) / 4
\]

\[
= (0, 9^2 + 0, 5^2 + 0, 4^2 + 1^2) / 4 = 0, 555
\]

Hieraus folgt, daß Landmarke $A$ bezogen auf die räumliche Lokation als Referenzobjekt zu präferieren ist. Wie aber später bei der vollständigen Beschreibung der Auswahl von Referenzobjekten gezeigt wird, sind die visuelle Salienz und der räumliche Anwendbarkeitsgrad einer Landmarke zu integrieren (vgl. Kapitel 4.2.6).

In diesem Kapitel sind die Konzepte vorgestellt worden, mit denen das räumliche Wissen über eine Situation repräsentiert. Bedingt durch Beschränkungen, die durch die Ressource Zeit und durch Speicherplatzrestriktionen vorgegeben sind, wird nur eine minimale Menge an räumlichen Relationen etabliert, die zur Handlungsfindung notwendig sind. Bezüglich einer Situation wird zuerst eine minimale räumliche Konfigurationsbeschreibung generiert, die die relevanten Straßenelemente zum Agenten und zueinander in räumliche Beziehung setzt. Eine minimale räumliche Konfigurationsbeschreibung enthält das minimale räumliche Wissen, welches notwendig ist, um einen Wegabschnitt adäquat zu beschreiben. Steht dem Agenten genügend Zeit zur Verfügung, kann die räumliche Konfigurationsbeschreibung durch Landmarkenwissen erweitert werden. Eine solche Repräsentation wird als erweiterte räumliche Konfigurationsbeschreibung bezeichnet.

Bevor die Verbindung zwischen räumlichem Wissen und sprachlichen Beschreibungen vorgestellt wird, diskutiert das folgende Kapitel die qualitativen Navigationsoperationen, die dem Agenten zur Eigenbewegung zur Verfügung stehen.

\footnote{Bezogen auf den vorliegenden Korpus hat sich diese Definition des Anwendbarkeitsgrades einer Landmarke als geeignet herausgestellt. Dies liegt u.a. an der Eigenschaft der Quadratfunktion, daß der relative Unterschied zwischen kleineren und größeren Werten vergrößert wird. Ferner besitzt die Funktion die gewünschten Eigenschaften streng monoton steigend und positiv zu sein. Weitere, dedizierte empirische Untersuchungen müßten zeigen, ob dieser Zusammenhang allgemein gilt.}
4.2. FUNKTIONALE BESCHREIBUNG

4.2.5 Funktionen des Navigationsmoduls


Abbildung 4.58: Ableitung der Lokation und Orientierung des PRS aus der Information des motorischen Moduls bezüglich Lokationen, die sich in einer Distanz s vor der Kreuzung und auf der Kreuzung befinden.

In der betrachteten stadtähnlichen Domäne können Bewegungen ausschließlich auf Pfadelementen ausgeführt werden. Sind der aktuelle, der nachfolgende Pfadabschnitt sowie der Verzweigungspunkt festgelegt und pfadbasierte Intentionen abgeleitet worden, so erhält das Navigationsmodul vom RI-M die Anweisung, entsprechend der topographischen Lagebeziehungen der Pfadelemente geeignete Navigationsfunktionen zu evaluieren. Eine translatorische Bewegung wird durch Angabe einer Geschwindigkeit \( v \) und einer Länge der zu bewältigenden Distanz \( s \) aktiviert (\( MOVE(v, s) \)). Die Aktivierung einer Rotation des gesamten Agenten\(^{55} \) erfolgt über die Angabe der Rotationsgeschwindigkeit \( r \) und dem Rotationswinkel \( \alpha \) (\( ROTATE(r, \alpha) \)).


In den vorangegangenen Kapiteln ist gezeigt worden, in welcher Weise raumbezogenes Wissen vom Agenten durch das Zusammenspiel mehrerer autonomer Module erfolgt. Als Ergebnis wird die räumliche Vorstellung, die der Agent über seine Umgebung besitzt, in einer räumlichen Konfigurationsbeschreibung repräsentiert. Im nächsten Kapitel wird die Umsetzung dieses raumbezogenen Wissens in sprachliche Beschreibungen diskutiert. Eine sprachliche Beschreibung ist jedoch ebenso wie andere Verhaltensweisen im Kontext einer Situation zu betrachten.

\(^{55}\)Dies steht im Gegensatz zur Etablierung eines sekundären Referenzsystems, bei der nur die Ausrichtung des perzeptuellen Apparates gedreht wird.
Abbildung 4.59: Schnappschüsse einer translatorischen Bewegung, die durch Evaluation von $\text{MOVE}(v, s)$ erzeugt wurde.

4.2.6 Funktionen des Sprachproduktionsmoduls

In diesem Kapitel wird die sprachbezogene generische Inhaltsstruktur vorgestellt, von der sich alle anderen Inhaltsstrukturen ableiten. Danach wird die operationale Vorgehensweise beschrieben, durch die der Inhalt einer inkrementellen Wegbeschreibung ausgewählt wird. Vor allem die Referenzierung von Landmarken wird eingehender diskutiert. Das Ergebnis dieser Funktion ist eine Oberflächenstruktur, die in sprachliche Beschreibungen (Terminalketten) transformiert wird. Abschließend wird darauf eingegangen, zu welchem Zeitpunkt eine Beschreibung gegeben wird.


4.2.6.1 Generische Inhaltsstruktur inkrementeller Wegbeschreibungen

Aufgrund der Schematisierung der kommunikativen Handlung einer inkrementellen Wegbeschreibung werden häufig einige Kategorien, insbesondere die Kategorie $\text{WER-K}$, die das Subjekt, und $\text{WAS-K}$, die die Aktion spezifiziert, i.e. elliptifiziert und durch den Kontext als bekannt vorausgesetzt. Dies beruht darauf, daß der Rezipient sich während einer Beschreibung nicht ändert und
4.2. **FUNKTIONALE BESCHREIBUNG**

Abbildung 4.60: Schnappschüsse einer rotationellen Bewegung, die durch Evaluation von \( \text{ROTA-TE}(r, \alpha) \) erzeugt wurde.


Wie die Analyse des Korpus ergibt (vgl. Kapitel 3), lassen sich die Beschreibungen inhaltsmäßig durch eine \emph{generische Inhaltsstruktur} zusammenfassen (siehe auch Abbildung 4.61), welche die Zusammenhänge zwischen Instanzen von Inhaltskategorien in einzelnen Beschreibungen repräsentiert. Von der generischen Inhaltsstruktur werden, entsprechend den zeitlichen und räumlichen Beschränkungen, \emph{situationsbezogene Inhaltsstrukturen} abgeleitet. Eine inkrementelle Wegbeschreibung besteht im allgemeinen Falle aus einem Ereignis (EVENT), in dem der Rezipient aufgefordert wird, eine Aktion (ACTION) auf einem Pfad (PATH) an einer bestimmten Lokation (\emph{PLACE}_{\text{dia}}) auszuführen.\footnote{Diese Notation einer Inhaltsbeschreibung entspricht der von Levelt (vgl. [Lev89]).} In Fällen eines langen Pfadabschnitts wird dem Hörer eine unterstützende, \emph{deklamative} Beschreibung gegeben, die indiziert, daß man sich immer noch auf dem richtigen Weg befindet. Die Beschreibung einer Aktion wird in einem \emph{imparativen} Beschreibungsmodus gegeben.

Inkrementelle Wegbeschreibungen unterscheiden sich vor allem in dem Teil, welcher eine auszuführende Aktion beschreibt. In diesem Teil ist eine Instanz aus WAS-K (\emph{GO}_{\text{phys}}), aus WANN-
Abbildung 4.61: Generische Inhaltsstruktur für inkrementelle Wegbeschreibungen und die Abfolge der Instantiierung


\[(41) \quad \text{„An der Kreuzung rechts ab.“}\]


### 4.2.6.2 Auswahl von Inhaltsstrukturen bezüglich einer Situation


---

\(^69\) Referenziert die Lokation, an dem sich der Agent momentan befindet.

\(^70\) Referenziert die Lokation, über die der Agent in den nächsten Pfadabschnitt gelangen möchte.

\(^71\) Zusätzlich sind Verben wie „einordnen“, die ebenfalls adverbiale Erweiterung erhalten können.
relevanten Kategorien instantiiert worden sind, wird die Oberflächenstruktur ausgewählt, deren Gewichtung, entsprechend den empirischen Untersuchungen (vgl. Tabelle 3.2 in Kapitel 3.2), maximal ist.

**Funktion 11: Generate-IR-Description**

\[
\text{Generate-IWB-Description}(T_{\text{guess}}, \text{listener} \_\text{props}, T_{\text{old}}) \\
\text{content-structure-set} := \text{Determine-Content-Structure-Set}(T_{\text{guess}}, \text{listener} \_\text{props}, T_{\text{old}}); \\
\text{transformation-set} := \text{Generate-WHEN-K}(\text{transformation-set}); \\
\text{transformation-set} := \text{Generate-WHERE-TO-K}(\text{transformation-set}); \\
\text{transformation-set} := \text{Generate-WHAT-K}(\text{transformation-set}); \\
\text{surface-structure} := \text{Select-Most-Specified}(\text{transformation-set}); \\
\text{return}(\text{surface-structure});
\]

Beruht eine Beschreibung auf der gesamten generischen Inhaltsstruktur, erfolgt die Transformation der Inhaltsstruktur des Aktionsteils in eine Oberflächenstruktur in vier Schritten. Vom Zeitrahmen \( T_{\text{guess}} \) ausgehend, wird in einem ersten Schritt eine natürlichsprachliche Beschreibung einer zeitleichen Referenz abgeleitet (WANN-K, 1. Schritt in Abbildung 4.61). Dieser Evaluation nachgeordnet ist die zur Instantierung von \( WOHIN-K \), welches die Orientierung einer Aktion indiziert (2. Schritt). Die zugrundeliegende Richtungsinformation wird von den pfadbasierten Intentionen abgeleitet. Die Etablierung einer geeigneten Instanz der \( WOHIN-K \) Kategorie initiiert die Auswahl einer \( WAS-K \) Instanz (3. Schritt). Entweder handelt es sich um einen Abbiegevorgang oder um eine Fortsetzung der Bewegung in Geradeausrichtung. Falls zeitleiche Beschränkungen es zulassen, der Rezipient es erwünscht und geeignetes Landmarkenwissen vorhanden ist, wird in einem vierten Schritt die Lokation, an der die Aktion stattfinden soll, genauer spezifiziert. Eine Lokation kann durch unterschiedlich viele Landmarken und deren Eigenschaften beschrieben werden, was durch eine Sequentialisierung einzelner \( WO-K \) Instanzen erreicht wird.\(^2\) In seltenen Fällen wird auch die aktuelle Lokation als Referenzpunkt eingeführt (PLACE\(_\text{from}\)). Die Transformation einer Instanz einer solchen Kategorie erfolgt analog zur \( WO-K \) Instanz, die den Verzweigungspunkt indiziert (PLACE\(_\text{via}\)).

### 4.2.6.2.1 Auswahl von situationsbezogenen Inhaltsstrukturen

Zeitliche Beschränkungen beziehen sich auf das Verhältnis zwischen dem Agenten, dem Rezipienten und der Umgebung. Alle Beschreibungen, die der Hörer aufgrund ihrer Komplexität nicht bis zum Erreichen des Verzweigungspunktes verstehen und umsetzen kann, gelten als nicht adäquat.\(^3\) Wie aus dem Korpus hervorgeht (siehe Analyse der zeitlichen Kategorien in Kapitel 3), lassen sich fünf Zeitrahmen unterscheiden, in denen die Probanden unterschiedlich strukturierte Beschreibungen gegeben haben (vgl. Kapitel 3, 4.2.1). Die Annahme bei der Bestimmung einer Beschreibung lautet, daß ein Sprecher immer versucht, in einer gegebenen Situation eine möglichst gute Beschreibung zu generieren. Hierzu dient die Heuristik, daß je genauer die Information, um so besser die jeweilige Beschreibung ist. Über das Zeitintervall \( T_{\text{guess}} \) werden a priori situationsabhängig Inhaltsstrukturen von der generischen Inhaltsstruktur abgeleitet.

Hierzu berechnet die Funktion \( \text{Determine-Content-Structure-Set} \) (vgl. Funktion 12), mit welchem Zeitrahmen der aktuelle Zeitrahmen \( T_{\text{guess}} \) übereinstimmt (\( \text{Match-Current-Time-Frame} \)). Nicht in jedem Fall wird dem Hörer eine maximal detaillierte Beschreibung gegeben. Ist der

\(^2\) Diese Iterationsmöglichkeit ist in der generischen Inhaltsstruktur durch einen Stern markiert: \( WO^* \).

\(^3\) Ein solches Konzept der \textit{Adäquathheit} gilt nicht für eine dialogische Situation. In einem Dialog kann der Rezipient im Zweifelsfalle nachfragen, bzw. sich rückerseichern.
Funktion 12: Determine-Content-Structure-Set

\[
\text{Determine-Content-Structure-Set}(T_{\text{guess}}, \text{listener}_\text{props}; T_{\text{old}}) \\
T_{\text{current}} := \text{Match-Current-Time-Frame}(T_{\text{guess}}); \\
T_{\text{listener}} := \text{Match-Listener-Time-Frame}(\text{listener}_\text{props}) \\
T_{\text{select}} := \min(T_{\text{current}}, T_{\text{listener}}, T_{\text{old}}); \\
\text{content-structure-set} := \text{Select-Structure-Schema}(T_{\text{select}});
\]

Hörer an einer möglichst detaillierten Beschreibung interessiert, so wird durch die Funktion 
\text{Match-Listener-Time-Frame} der Wert von \(T_{\text{listener}}\) auf \(\infty\) gesetzt, so daß der maximal mögliche Zeitrahmen durch \(T_{\text{current}}\) vorgegeben wird. Ist zu der aktuellen Situation bereits eine Beschreibung im Zeitrahmen \(T_{\text{old}}\) gegeben worden, so ist dies für die nachfolgende Beschreibung zu beachten. Eine nachfolgende Beschreibung kann dann erst im übernächsten Zeitrahmen \((T_{\text{old}} + 2)\) gegeben werden. Ist \(T_{\text{listener}}\) ungleich \(\infty\), so ist der maximale Zeitrahmen \(T_{\text{select}}\), der für die Beschreibung zur Verfügung steht, durch das Minimum über die Zeitrahmen \(T_{\text{guess}}, T_{\text{listener}}\) und \(T_{\text{old}} + 2\) bestimmt. Dieses spiegelt wider, daß der Sprecher den Rezipienten nicht unstrukturiert mit Information überfordert, sondern eine Grobstruktur verwendet, die nur zu bestimmten Zeitrahmen Beschreibungen ermöglicht. Das Zeitintervall \(T_{\text{guess}}\) gibt vor, zu welchem Zeitrahmen \(T_1\) die nächste Beschreibung potentiell generiert werden kann (vgl. Kapitel 4.2.1). Jedem der fünf Zeitrahmen lassen sich Ableitungen der generischen Inhaltsstruktur zuordnen, die unterschiedliche Kombinationen von \(W^\ast\)-Kategorien enthalten (vgl. Abbildung 4.62). Je näher dieser Zeitrahmen vor dem Erreichen des Verzweigungspunktes ist, um so kürzer ist die resultierende Beschreibung. Von den beiden Zeitrahmen \(T_{\text{current}}\) und \(T_{\text{listener}}\) wird derjenige, der am nächsten zum Verzweigungspunkt ist, ausgewählt. Bezüglich dieses Zeitrahmens bestimmt sich die Menge der möglichen Inhaltsstrukturen (siehe Tabelle 3.2 in Kapitel 3.2 und Abbildung 4.62). Tabelle 3.2 zeigt eine Zuordnung der einzelnen Kategorien zu Zeitrahmen, wobei jeder einzelnen Zeile eine Konjunktion von Inhaltskategorien zugeordnet ist, die in Abbildung 4.62 graphisch illustriert ist. Die Gewichtung gibt an, mit welcher Häufigkeit eine situationsabhängige Inhaltsbeschreibung bezüglich eines Zeitrahmens verwendet wird (vgl. Kapitel 3).

Wenn die Menge der Inhaltsstrukturen für die aktuelle Situation ausgewählt worden ist, werden die Kategorien der einzelnen Inhaltsstrukturen in Oberflächenstrukturen transformiert. Der Rezipient ist in allen Fällen konstant und wird optional durch eine geeignete Nominalphrase gefüllt. Danach werden zeitbezogene, gefolgt von intentionsbezogenen und abschließend lokationsbezogenen Kategorien im Wissen des Agenten über die Situation verankert. Durch diese Schritte erfolgt die Transformation von Inhaltsstrukturen in Oberflächenstrukturen. Bei der Instantiierung einer Inhaltskategorie wird die Wortwahl getroffen und dadurch, entsprechend Levelt's Hypothese ([Lev89]), die Phrasenstruktur der Instanz über grammatisch Rablige Regeln an die gesamte Struktur vererbt.

4.2.6.2.2 Zeitliche Referenzierung Zeitliche Referenzen auf sprachlicher Ebene korrespondieren eng mit solchen auf der Steuerungsebene. Über die Größe von \(T_{\text{guess}}\) entscheidet die Funktion \text{Generate-WHEN-Category} (vgl. Funktion 13), ob und welches Lexem eine geeignete Beschreibung der zeitlichen Zusammenhänge zwischen der Bewegung des Agenten und der Situation repräsentiert.

Wie der Tabelle 3.2 in Kapitel 3.2 entnommen werden kann, enthalten verschiedene Inhaltsstrukturen Kategorien des Typ \text{WANN-K}. Geeignete Lexeme werden durch die Aktivatoren @now, @soon, @later ausgewählt, die zu einer Adverbalphrase transformiert werden. Dabei wird die
adverbiale Ergänzung des Verbs um die der zeitlichen Referenz erweitert. Die Funktion *Lexica-lize* ersetzt in der Inhaltsstruktur einen TIME-Marker, der durch die ersten Parameter gegeben ist, durch eine Phrasenstruktur, die dem Lexem in der zweiten Position zugeordnet ist.

4.2.6.2.3 Instantiierung einer Richtungs- und Aktionsreferenz  Die Bedeutung einer WOHN-K Kategorie ist es, die Orientierung der am nächsten Verzweigungspunkt auszuführenden Aktion zu indizieren (vgl. Funktion 14). Entsprechend der pfadbasierten Intention mit dem höchsten Anwendbarkeitsgrad AG_{int} bzw. AG_{kpi} wird eine geeignete Adverbialphrase ausgewählt, die die Instanz der WOHN-K Kategorie ersetzt. Die pfadbasierte Intention hat Auswirkungen auf die Transformation der WAS-K Instanz. Handelt es sich um einen Abbiegevorgang, d.h., daß die Intention $\$turn-left\$ oder $\$turn-right\$ enthält, wird durchweg eine Verbphrase etabliert, welche eine solche Aktion beschreibt (vgl. Funktion 15).

4.2.6.2.4 Integration von räumlichem Landmarken- und Routenwissen  Der Hauptbestandteil einer inkrementellen Wegbeschreibung bezieht sich auf eine elaborierte Verwendung
Funktion 13: Generate-WHEN-Category

\[
\text{Generate-WHEN-Category} (\text{transformation-set})
\]
\[
\text{Case } T_{\text{select}} \text{ is equal}
\]
\[
T_5: \text{Lexicalize(TIME, } \text{@now)}
\]
\[
T_4: \text{Lexicalize(TIME, } \text{@soon)}
\]
\[
T_3: \text{Lexicalize(TIME, } \text{@later)}
\]

Funktion 14: Generate-WHERE-TO-Category

\[
\text{Generate-WHERE-TO-Category} (\text{transformation-set})
\]
\[
\text{If } \# \text{right-of}(\text{ep, } p_2) \text{ then}
\]
\[
\text{Lexicalize(PLACE-TO, } \# \text{right#, nil)}
\]
\[
\text{if } \# \text{left-of}(\text{ep, } p_2) \text{ then}
\]
\[
\text{Lexicalize(PLACE-TO, } \# \text{left#, nil)}
\]
\[
\text{if } \# \text{behind}(\text{ep, } p_2) \text{ then}
\]
\[
\text{Lexicalize(PLACE-TO, } \# \text{straight#, nil)}
\]


4.2.6.2.4.1 Referenzierung von Landmarken  Eine Landmarke wird dann in die Beschreibung integriert, wenn sie sich in einer geeigneten räumlichen Lokation zum aktuellen Pfadabschnitt \( p_{\text{current}} \) zum nächsten Pfadabschnitt \( p_{\text{next}} \) und zum Verzweigungspunkt \( e_p \) befindet (vgl. Funktion 18). Ob eine Landmarke als Referenz für die Beschreibung einer Aktion verwendbar ist, hängt vom räumlichen Anwendbarkeitsgrad einer Landmarke (RAG) relativ zu den Pfadedelementen als auch von seiner visuellen Salienz (VS) ab. Allgemein läßt sich der Anwendbarkeitsgrad (AG) durch eine Vermischung dieser beiden Eigenschaften einer Landmarke durch zwei Funktionen \( F \) und \( G \) angeben (vgl. Gleichung 42).

\[
AG(\text{lm}) = F(\text{RAG(\text{lm})}) \otimes G(\text{VS(\text{lm})})
\]

Aufgrund der Einfachheit ist eine Linearkombination eine erste Approximation desOperators \( \otimes \). Bezogen auf das Experiment zur Untersuchung der visuellen Salienzen hat sich die Identitätsfunktion als geeignete Modellierung für \( F \) herausgestellt (vgl. Kapitel 3.3). Die Funktion \( G \) bezieht sich auf die visuelle Salienz (VS(\text{lm})) einer Landmarke (vgl. Kapitel 4.2.3). Hierüber ist repräsentiert, ob die Landmarke visuell von anderen Landmarken unterscheidbar ist und somit maximal disambiguierend wirkt (vgl. Kapitel 4.2.3). Visuelle Salienzen haben einen stärkeren Einfluß auf die Anwendbarkeit einer Landmarke, so daß die Euler-Funktion eine geeignete Charakteristik bietet. Daran ergibt sich für den Anwendbarkeitsgrad einer Landmarke eine Linearkombination aus räumlicher Anwendbarkeit und visueller Salienz (vgl. Gleichung 43).

\[
AG(\text{lm}) = RAG(\text{lm}) + e^{VS(\text{lm})}
\]
4.2. FUNKTIONALE BESCHREIBUNG

Funktion 15: Generate-WHAT-Category

*Generate-WHAT-Category*(transformation-set)
   If #right-of#(ep, p₂) ∨ #left-of#(ep, p₂) then
      Lexicalize(GO, turn, nil)
   If #behind#(ep, p₂) then
      Lexicalize(GO, go, nil)

Funktion 16: Generate-WHERE-VIA-Category

*Generate-WHERE-VIA-Category*(transformation-set)
   lm-set := Select-AppI-LM-Set;
   if lm-set ≠ nil then
      Descr-Location-By-LM(lm-set)
   else
      Descr-Location-By-Path-Elements

Nicht in jedem Fall reicht der maximale Anwendbarkeitsgrad einer Landmarke aus, um sie als Referenzobjekt verwenden zu können. Besitzt eine Landmarke in einer Situation zwar den höchsten Anwendbarkeitsgrad AG, aber ist dieser zu gering, so wird die Landmarke nicht als Referenzobjekt verwendet.\(^{74}\) Als Rückgabewert liefert in diesem Fall die Funktion AppI-Degree den Wert $$-\infty$$. Im anderen Fall berechnet sie den Anwendbarkeitsgrad der Landmarke bezüglich der aktuellen Situation (vgl. Funktion 18).

In Abbildung 4.63 ist ein Beispiel gezeigt, in der die Landmarke A einen maximalen Anwendbarkeitsgrad besitzt, der aber nicht ausreicht, um A aufgrund visueller Merkmale als Referenzobjekt verwenden zu können.


Bei der Instantiierung der *WO-K* Kategorie lassen sich drei Fälle unterscheiden. Liegt eine Abbiegeintention vor, so kann der Verzweigungspunkt und der nachfolgende Pfadabschnitt entweder „hinter“ oder „vor“ der zu referenzierenden Landmarke liegen. Der dritte Fall betrifft

Funktion 17: Select-Appl-LM-Set

\[
\begin{align*}
\text{Select-Appl-LM-Set} & \quad \text{max-appl-degree} := 0; \\
& \quad \text{for all lm \in Salient-Landmarks(rkb) do} \\
& \quad \quad \text{ad} := \text{Appl-Degree(lm)}; \\
& \quad \quad \text{if ag >= max-appl-degree then} \\
& \quad \quad \quad \text{integrate-lm-set} := \text{union(lm, integrate-lm-set)} \\
& \quad \quad \text{if ag > max-appl-degree then} \\
& \quad \quad \quad \text{max-appl-degree} := \text{ad}; \\
& \quad \quad \quad \text{integrate-lm-set} := \{\text{lm}\} \\
& \quad \text{return(integrate-lm-set)};
\end{align*}
\]

Funktion 18: Appl-Degree(lm)

\[
\begin{align*}
\text{Appl-Degree(lm)} & \quad \text{if AG(lm) >= appl-degree-threshold then} \\
& \quad \quad \text{return(appl-degree)} \\
& \quad \text{else} \\
& \quad \quad \text{return(-\infty)}
\end{align*}
\]

eine Intention der Fortbewegung \textit{geradeaus} in Frontalrichtung, wobei keine Landmarken als Referenzobjekte verwendet werden (vgl. Abbildung 4.64b).


Die Anwendbarkeit (AGP) einer räumlichen Präposition (sp-prep) wird durch das arithmetische Mittel der quadrierten Anwendbarkeitsgrade der räumlichen Relationen der assoziierten Landmarke bestimmt (vgl. Definition 44).

\[
\text{(44) \quad AGP(sp-prep) = } \frac{1}{n} \sum_{i=1}^{n} rrel_i(PAGM(prep, lm))
\]

Der Anwendbarkeitsgrad einer räumlichen Präposition mit dem höchsten Wert wird zur Beschreibung der räumlichen Lage des Gebietes, in dem die Aktion auszuführen ist, und der referenzierenden Landmarke verwendet. Die Bestimmung einer AGP erfolgt durch Vergleich der intrinsischen Relationen der räumlichen Konfigurationsbeschreibung bezüglich der Landmarke mit den geforderten Relationen der assoziierten PAGM. Die Auswertung erfolgt in zwei Schritten. Zuerst wird überprüft, ob alle vom PAGM geforderten räumlichen Relationen zwischen der Landmarke und den Pfadelementen bestehen. Ist dies nicht der Fall, scheitert die Auswertung. Bestehend alle Relationen, so werden sie entsprechend der Funktion \textit{AGP(sp-prep)} ausgewertet. Die räumliche
Abbildung 4.63: Beispiel für eine Situation, in der eine Landmarke A eine maximale visuelle Salienz besitzt, die aber relativ zu den anderen Objekten für eine Referenzierung zu gering ist.

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>C</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMA</td>
<td>#right-of# = 0,8</td>
<td>#right-of# = 0,7</td>
<td>#right-of# = 0,9</td>
</tr>
<tr>
<td></td>
<td>#in-front-of# = 0,6</td>
<td>#behind# = 0,7</td>
<td></td>
</tr>
</tbody>
</table>

Tabelle 4.4: Intrinsische Relationen und Anwendbarkeitsgrade zwischen der Landmarke A und relevanten Pfadelementen

Präposition mit dem maximalen Anwendbarkeitsgrad wird letztendlich ausgewählt.

In Abbildung 4.64a ist eine Beispielsituation, in der A die zu referenzierende Landmarke ist. Geht man davon aus, daß zwischen A und den beiden Pfadelementen C und S2 alle notwendigen räumlichen Relationen wie in Tabelle 4.4 bestehen, so ergibt sich ein maximaler PAGM bezüglich einer pfadbasierten Intention $\$turn-left$ für die räumliche Präposition „hinter“ (vgl. Berechnung 45). Die Auswertung der PAGM(vor) scheitert aufgrund der fehlenden #in-front-of# Relation zur S2 und C.

\[(45)\quad \text{PAGM(hinter, A) = (0,9^2 + 0,7^2 + 0,7^2)} / 3 = 0,45\]

Über eine Funktion Det-Max-PAGM wird die räumliche Präposition spatial-prep mit dem maximalen Anwendbarkeitsgrad AGP bestimmt. Liegt der AGP dieser räumlichen Präposition über einen zugeordneten Schwellwert Threshold(spatial-prep), so wird diese Präposition dazu verwendet, die Lokation der Aktion bezüglich der referenzierenden Landmarke zu beschreiben (vgl. Funktion 19)


Die Beschreibung einer Landmarke selbst wird durch die Funktion Descr-LM bestimmt (vgl. Funktion 21). Sie bestimmt zu einer gegebenen Landmarke geeignete Merkmale, die es dem Rezipienten erlauben, diese Landmarke ohne Ambiguitäten zu identifizieren.\footnote{Im Rahmen dieser Arbeit handelt es sich dabei ausschließlich um physikalisch-visuelle Merkmale.} Über die Men-
Abbildung 4.64: Drei Grundfälle, die bei der Referenzierung durch eine Landmarke auftreten können

Funktion 19: Descr-Location-By-LM

Descr-Location-By-LM(lm-set)
For all lm ∈ lm-set
  spatial-prep := Det-Max-PAGM(lm)
  if AGP(spatial-prep) ≥ Threshold(spatial-prep) then
    Lexicalize(PLACE-VIA, spatial-prep, (Descr-LM(lm-set)))
    return (t);
  else
    return (false);

Abbildung 4.65: Anwendbarkeitsmuster PAGM der räumlichen Präpositionen „vor“, „hinter“ und „zwischen“ bei unterschiedlichen pfadbezogenen Intentionen

Die Landmarke und die Merkmale werden durch eine Adjektivphrase lexikalisiert.

### 4.2.6.2.4.2 Referenzierung von Pfadelementen

Existieren keine Landmarken mit einem ausreichend hohen Anwendbarkeitsgrad, werden Pfadelemente dazu verwendet, das Gebiet der auszuführenden Aktion genauer zu spezifizieren. Eine Analyse des Korpus ergibt, daß die Beschreibung eines Gebietes der auszuführenden Aktion, basierend auf Pfadelementen, grundsätzlich durch eine Referenzierung des Verzweigungspunktes unter Verwendung der räumlichen Präposition „an“ erfolgt.

Besitzt der Verzweigungspunkt visuell-saliente Merkmale, werden diese zur genaueren Beschreibung verwendet (*Descr-Path-Element*, analog zu *Descr-LM*). In dieser werden auffällige Eigenschaften des Verzweigungspunktes extrahiert und durch die Funktion *Lexicalize* in Oberflächenstrukturen transformiert.\(^7\)

### 4.2.6.2.4.3 Lexikalisation und grammatikalische Regeln inkrementeller Wegbeschreibungen

Der Zugriff auf das Lexikon, die Wortwahl und der Aufbau einer syntaktischen
Funktion 20: Spatially-Consistent

\[
\text{Spatially-Consistent}(\text{lm}, \ p_1, \ p_2, \ \epsilon_p) \\
\text{if spatially-near(\text{lm}, \ p_2) then} \\
\quad \text{if Type(\text{SpRel}(\text{lm}, \ p_2)) == Type(\text{SpRel}(p_2, \ \epsilon_p)) then} \\
\quad \quad \text{return}(t); \\
\quad \text{if \#behind(\epsilon_p, \ p_2) \geq \text{Threshold(\#behind)}) then} \\
\quad \quad \text{return}(t); \\
\text{else:} \\
\quad \text{return}(\text{nil});
\]

Funktion 21: Descr-LM

\[
\text{Descr-LM}(\text{lm-set}) \\
\text{For all l \in \text{All-Selected-LMs \land l not \in lm-set do} \\
\quad \text{f-set}_l := \text{Salient-Features}(l); \\
\quad \text{f-set} := \text{union}(f-set, \ f-set_l) \\
\quad \text{f-set}_\text{lm} := \text{Salient-Features}(\text{lm}); \\
\quad \text{distinct-f-set} := \text{Select-Distinctive-Features}(\text{f-set}_\text{lm}, \ f-set); \\
\quad \text{Lexicalize}(\text{PLACE}_{\text{via}}, \ \text{first distinct-f-set}); \ (\text{rest distinct-f-set}, \ \text{lm})
\]

Teilphrase erfolgt durch die Funktion \text{Lexicalize} (vgl. Funktion 24). In \text{Lexicalize} sind die grammatischen Regeln kodiert, die durch Analyse des Korpus abgeleitet worden sind. Tabelle 4.5 enthält die aus dem Korpus abgeleiteten grammatischen Regeln inkrementeller Wegbeschreibungen. Als ersten Parameter übernimmt \text{Lexicalize} einen Terminalknoten einer Inhaltsstruktur, welcher entsprechend dem zweiten und dritten Parameter transformiert werden soll. Der zweite Parameter trägt den Wert des Phrasenkopfes, entsprechend dessen Typs die geeignete Phrasenstruktur ausgewählt wird. Der dritte Parameter enthält den Teil, der dieser Phrase untergeordnet ist. Durch rekursiven Aufruf wird eine Ersetzung dieses Teils erreicht. Der durch den dritten Parameter entstandene Teilbaum der Oberflächenstruktur wird an die geeignete Stelle der Phrase eingehängt. Strukturelle Merkmale werden entsprechend unifikationsgrammatikalische Techniken vererbt (\text{}). Eine genauere Analyse dieser Vererbungsmechanismen liegt außerhalb des Fokus dieser Arbeit (für verschiedene Ansätze siehe u.a. [Kay79], [Usz86]).\text{78} Der Rückgabewert der Funktion ist eine instanzierte Phrasenstruktur \text{}.


Gibt es nach Evaluation dieser Funktion alternative Oberflächenstrukturen zur Beschreibung der Situation, wird unter diesen eine solche ausgewählt, die einerseits vollständig instantiiert ist und die andererseits eine maximale Gewichtung entsprechend Tabelle 3.2 in Kapitel 3.2 besitzt (Funktion \text{Select-Most-Specified}).

\text{78}In der Implementation wird u.a. ein TAG-Generator (vgl. [Jos85], [HFS91], [KF95]) verwendet (vgl. [Pau05]).
\text{79}Nach der \text{X}-Grammatik von Jackendoff (vgl. [Jac72]) werden Regeln nicht beliebig rekursiv, sondern nur entsprechend einer fest vorgegebenen Tiefe evaluiert. Eine solche Tiefe ist für jede Phrase durch die Anzahl der Hochkommaten angegeben.
4.2. **FUNKTIONALE BESCHREIBUNG**

**Funktion 22: Select-Distinctive-Features**

*Select-Distinctive-Features* (*f-set*<sub>IM</sub>, *f-set*)

\[ \text{dist}-f\text{-set}<sub>IM</sub> := \text{nil} \]

For all feature \( f \) in *f-set*<sub>IM</sub> do

If Feature-Distance(*feature*, *f*) \( \geq \) Threshold(*F-Type*(*feature*)) then

\[ \text{dist}-f\text{-set}<sub>IM</sub> := \text{union} \left( \text{dist}-f\text{-set}<sub>IM</sub> \right, *feature*) \]

f-counter++;

If f-counter \( \geq \) max-distinct-features then

return (dist-f-set<sub>IM</sub>)

done

return (dist-f-set<sub>IM</sub>)

**Funktion 23: Descr-Location-By-Path-Elements**

*Descr-Location-By-Path-Elements*

\[
\text{Lexicalize}\left( \text{PLACE}_{\text{via}} \cdot \text{an'} \cdot (\text{Descr-Path-Element}(ep)) \right)
\]

4.2.6.3 **Beispiele der Verwendung räumlicher Konfigurationsbeschreibungen zur Generierung inkrementeller Wegbeschreibungen**


4.2.6.4 **Der Präsentationsspeicher**


Bevor eine neue Beschreibung im Präsentationspeicher übernommen wird, wird der Speicher initialisiert um die bisherige Beschreibung zu deaktivieren (Initialize-Presentation-Buffer). Die Repräsentation der Beschreibung enthält die Information über den Zeitrahmen ts. Diese temporale Information dient dem Präsentationspeicher dazu, den Zeitpunkt zu bestimmen, an dem die Beschreibung in der Situation gegeben werden soll.

Die Notwendigkeit für die Autonomie des mit der Funktion *Activate-Presentation-Buffer* assoziierten Prozesses ist stark durch die Ressourcenadaptivität des Gesamtmodells vorgegeben (vgl.
### Tabelle 4.5: Grammatikalische Regeln inkrementeller Wegbeschreibungen

4.3 Implementation des inkrementellen Wegbeschreibungsagenten MOSES

Das in Kapitel 4.2 beschriebene funktionale Modell ist die Grundlage für die Implementation des inkrementellen Wegbeschreibungsagenten MOSES. MOSES ist mittels einer Multiagentenarchitektur realisiert, in der jedes einzelne Modul durch einen Agenten modelliert wird. Dabei sind Sequentialisierungsbeziehungen zwischen Modulen durch eine asynchrone bzw. synchrone Aktivierung der Module repräsentiert. Als Modell einer stadtähnlichen Umgebung wird ein dreidimensionales, geometrisches Modell des Campus der Universität des Saarlandes verwendet.

Die Implementation von MOSES besteht aus drei Teilsystemen. Das erste Teilsystem, MOSES-VIEW, ist die Implementation des visuellen Objektauswahlmoduls, sowie der Animationen, die es dem Agenten erlauben, sich in der Umgebung zu bewegen (vgl. Kapitel 4.1.3, 4.2.3). Das zweite Modul MOSES-PATH ist die Implementation des Pfadsuchmoduls (vgl. Kapitel 4.1.2, 4.2.2). Letztlich wird das raumbezogene Wissen im Teilsystem MOSES-SPEAK etabliert und in sprachliche Beschreibungen umgesetzt (vgl. Kapitel 4.1.4, 4.2.4, 4.1.6, 4.2.6).

Das System MOSES ist vollständig in der objektorientierten Programmiersprache LISP (Lucid-CLOS) implementiert worden. Die grafischen Oberflächen basieren auf dem Com-

---

Funktion 24: Lexicalize

\[
\text{Lexicalize(c-leaf, s-node, rec-node)} \\
\text{lex := Select-Lexem(s-node);} \\
\text{Case c-leaf is equal to} \\
\text{TIME:} \\
\text{Case Type(s-node) is equal} \\
\text{temporal-ref: p-type := AdvP;} \\
\text{GO:} \\
\text{Case Type(s-node) is equal} \\
\text{go: p-type := VP;} \\
\text{PLACE-FROM:} \\
\text{Case Type(s-node) is equal} \\
\text{deictic-ref: p-type := AdvP;} \\
\text{PLACE-VIA:} \\
\text{Case Type(s-node) is equal} \\
\text{sp-rel: p-type := PP;} \\
\text{iota: p-type := NP;} \\
\text{feature: p-type := AP;} \\
\text{object: p-type := NP;} \\
\text{PLACE-TO:} \\
\text{Case Type(s-node) is equal} \\
\text{sp-rel: p-type := AdvP;} \\
\text{feature: p-type := AP;} \\
\text{object: p-type := NP;} \\
\text{nil:} \\
\text{return(nil)} \\
\text{surfstruct := Generate-Surf-Struct(p-type, lex,} \\
\text{Lexicalize(PLACE-VIA, (first rec-node), (rest rec-node)))} \\
\text{return(surfstruct)}
\]
Abbildung 4.66: Verwendung räumlichen Wissens in inkrementellen Wegbeschreibungen

**mon Lisp Interface Management System (CLIM1.0).** Das Gesamtsystem MOSES besitzt eine Quellcodegröße von 1,3 Mega Byte. Dabei entfallen auf:

1. **MOSES-PATH:** 374 KByte,
2. **MOSES-VIEW:** 390 KByte,
3. **MOSES-SPEAK:** 567 KByte.

Das System MOSES agiert nahezu in Echtzeit, wobei ein großer Anteil des Berechnungsaufwandes durch das Subsystem MOSES-VIEW in Anspruch genommen wird. In Abbildung 4.67 ist die interaktive Bedienungsoberfläche gezeigt, über die ein Benutzer Anfragen in Form einer mausgestützten Auswahl von Start- und Ziellokationen stellen kann. Eine Anpassung an den Benutzer erfolgt über die Auswahl von Alternativen über ein Menü, welches über die Schaltfläche STATUS aktiviert werden kann (vgl. [Pau95]). Vor allem darüber wird das Verhalten von MOSES durch die Geschwindigkeit und den Bewegungstyp verändert. Als Ergebnis wird die Bewegung des Agenten auf einer Straßenkarte und die passende perspektivische Sicht auf die
Funktion 25: Activate-Presentation-Buffer

\[
\text{Activate-Presentation-Buffer(description)} \\
\text{Initialize-Presentation-Buffer;} \\
ts_0 := \text{Det-Time-Slot(description);} \\
\text{Present-Description(description, ts);} \\
\text{Initialize-Presentation-Buffer;}
\]

Abbildung 4.67: Bedienungsoberfläche von MOSES


durch erhält ein Agent Information von einem anderen Agenten, wird ein Agent durch eine Anfrage aktiviert bzw. wird eine Antwort auf eine Anfrage zurückgesendet. Entsprechend den Sequentialisierungsbedingungen ist der Modus vorgegeben, in der eine *ask* Nachricht versendet werden kann. Bestehst zwischen den Modulen, welche jeweils durch einen Agenten instantiiert werden, eine sequentielle Beziehung, so sind nur synchrone *ask* Nachrichten möglich. In diesem Fall wird der sendende Agent solange in einen Wartezustand versetzt, bis auf die Anfrage eine *reply* Nachricht des empfangenden Agenten gesendet wird.

Ein Beispiel für die Interaktion zwischen Agenten ist in Abbildung 4.6 gezeigt, die einen Ausschnitt eines Protokolls zur Generierung einer inkrementellen Wegbeschreibung repräsentiert.

```
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-128 and Content (RIM)
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-128 and Content NIL
AGENT-2 sends ASYNCH ASK to (AGENT-6) with Job-id JOB-129 and Content (WAKE-NAVIGATION (QUOTE (:SP056 :SP058 :SP025 :SP027 ; :SP026))
AGENT-2 sends ASYNCH ASK to (AGENT-4) with Job-id JOB-130 and Content (NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-130 and Content NIL
AGENT-6 sends SYNCH ASK to (AGENT-7) with Job-id JOB-131 and Content (V3::CLIM::GET-SALIENT-OBJECT)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and Content :KRISTALLOGRAPHIE
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-132 and Content (RIM)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and Content :KRISTALLOGRAPHIE
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-4) with Job-id JOB-133 and Content (NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-133 and Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-4) with Job-id JOB-134 and Content (NLP)
AGENT-4 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-134 and Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-135 and Content (RIM)
AGENT-3 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-135 and Content NIL
AGENT-2 sends SYNCH ASK to (AGENT-6) with Job-id JOB-136 and Content (WAKE-ANIMATION (QUOTE (:SP072))
AGENT-6 sends SYNCH ASK to (AGENT-7) with Job-id JOB-137 and Content (V3::CLIM::GET-SALIENT-OBJECT)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and Content :CHEMIE
AGENT-2 sends SYNCH ASK to (AGENT-3) with Job-id JOB-138 and Content (RIM)
AGENT-5 sends ASYNCH REPLY to (AGENT-2) with Job-id JOB-139 and Content :CHEMIE
```

Tabelle 4.6: Beispiel für eine Nachrichtensequenz zwischen Agenten in MOSES
4.3. IMPLEMENTATION

4.3.1 Prozesse der inkrementellen Pfadsuche

Die Implementation des Pfadmoduls basiert auf einem inkrementellen, heuristischen Suchalgorithmus, ähnlich dem RTA*-Algorithmus von Korf (vgl. [Kor90], [RW91]). Dieses Modul bestimmt lokal optimale Pfade (vgl. Kapitel 4.2.2). Die Repräsentation des Suchraums unterteilt sich in Knoten- und Kantenrepräsentationen, wenden im einzelnen strukturspezifische Merkmale zugeordnet sind.


Abbildung 4.68: Oberfläche des Systems MOSES-PATH mit eingekleinterem Korridor

Das Verhalten des inkrementellen Pfadsuchprozesses läßt sich über zahlreiche Parameter, wie den Öffnungswinkel des Korridors, die Größe der Vorausschau sowie die Gewichtung von Kantenarten modifizieren.\textsuperscript{92} Alle Parameter können zur Laufzeit verändert werden, so daß eine hohe Flexibilität erreicht wird. Über Trace-Ausgaben läßt sich das Verhalten des Pfadprozesses detailliert überprüfen.


\textsuperscript{92} Für eine detaillierte Diskussion vgl. [AK96].


Unter der Annahme, daß kognitive Prozesse eine globale Optimalität zugunsten lokaler Ökonomie vernachlässigen, ist die Verwendung des inkrementellen Ansatzes unter Verwendung einer
4.3. IMPLEMENTATION

Abbildung 4.70: Beispiel einer Pfadsuche mit einer Vorausschau von einem Pfadabschnitt

begrenzten Vorausschau vorzuziehen. Die Inkrementalität der Berechnung hält die Komplexität einer Bestimmung eines Pfadabschnitts konstant. Dabei verhält sich die Komplexität proportional zum Produkt aus maximalen Verzweigungsgrad und der Länge des *look-forward*. Die gesamte Komplexität eines Pfades berechnet sich aus der Summe der Einzelevaluationen.\(^{83}\)

4.3.2 Prozesse zur visuellen Objektauswahl

Die Implementation des Objektauswahl-Moduls MOSES-VIEW\(^{84}\) basiert auf den in Kapitel 4.2.3 vorgestellten Funktionen (vgl. Abbildung 4.72).


Der Blickwinkel, aus dem eine perspektivische Sicht auf das geometrische Modell des Universitäts-Campus betrachtet wird, hängt von der Lokation des Agenten (eye point), sowie seinem Blickpunkt (aim point) ab. Die Funktionen für die Modellierung des geometrischen Modells basieren auf dem System GEOMETRIX (vgl. [Her96]), sowie die Funktionen zur Bestimmung von animierten perspektivischen Sichten des Agenten auf dem System GEODISPLAY (vgl. [Müll95]). Das geometrische, dreidimensionale Modell ist durch Punkte, Kanten und Polygonzüge

---

\(^{83}\) Der hier vorgestellte inkrementelle, heuristische Suchalgorithmus ist in verschiedenen Domänen, wie beispielsweise auf einem Ausschnitt des deutschen Autobahnnetzes, getestet worden (vgl. [AK96]).

\(^ {84}\) Dieses Modul ist von Jörg Baus realisiert worden (vgl. [MBP95], [Bau95]).

\(^{85}\) Die Größe der Ellipsen ist durch die Größe des Objektes vorgegeben und besitzt aus diesem Grund keine weitere Bedeutung.
Abbildung 4.71: Beispiel einer Pfadsuche mit einer Vorausschau von zehn Pfadabschnitten

definiert, die zu komplexeren Objekten zusammengesetzt werden. In der vorliegenden Implementation sind Gebäude durch Quadere approximiert. (vgl. Tabelle 4.7).

Tabelle 4.7: Geometrische Definition des Objekts „Informatik“

In der vorliegenden Implementation von MOSES ist das zugrundeliegende Modell eine geometrische Repräsentation des Campus der Universität des Saarlandes. Die Grundfläche besteht aus einer Menge von planaren Polygonzügen. Weitere Elemente sind Pfade, Entscheidungspunkte
Abbildung 4.72: Bedienungsoberfläche des Subsystems MOSES-VIEW

(Kreuzungen, Plätze und Kreisverkehre) und Landmarken. Basierend auf dem Konzept eines dreidimensionalen Objektes sind dem Konzept einer Landmarke (landmark) weitere Eigenschaften wie ein Name, eine Farbe sowie eine Koordinate des Schwerpunkts zugeordnet (vgl. Tabelle 5).

Beispiel 5: Beispiel einer Instanz einer Landmarke


Beispiel 6: Beispiel einer Instanz eines Pfadabschnitts

Entscheidungspunkte (street-crossing) werden wie Pfadabschnitte als Rechteck definiert. An einen Entscheidungspunkt können jedoch eine beliebige Anzahl von Pfadabschnitten angrenzen (seg-1,..., seg-n) (vgl. Tabelle 7). Grundlage für die Berechnung visueller Salienz von Objekten ist die Bestimmung der Sichtbarkeit einzelner Objekte. Objektbezogene Flächen bezüglich des aktuellen eye point und aim
Beispiel 7: Beispiel einer Instanz eines Entscheidungspunktes

Eine mögliche Adaption von MOSES ist es, die Größe der Zellen ressourcenabhängig festzulegen. Je weniger Zeit zur Verfügung steht, um so geringer ist die Auflösung auf der Rasterene. Dies hat Auswirkungen auf die Berechnung visueller Salienzen und dadurch auf die Genauigkeit der Analyse der gesamten Situation. Steht mehr Zeit zur Verfügung kann die Situation besser, d.h. mit höherer Auflösung, analysiert werden.

Da in der vorliegenden Implementierung von einer homogenen Verteilung rekursiver Merkmale bezüglich eines Objekts ausgegangen wird, sind diese auf Rasterene und mitteln nicht mehr neu bestimmt werden.

Einer nebeneiligen Implementierung steht kein prinzipielles Hindernis im Wege. Da aber die Implementation von MOSES-VIEW für die Evaluation auf einem herkömmlichen sequentiellen Rechner geschrieben ist, ist die Sequentiellität inhärent.

Andere Formen der Foki, wie z.B. kreisförmige, sind denkbar, erfordern jedoch bezüglich einer Rasterrepräsentation einen erhöhten Berechnungsauflauf.

In der aktuellen Implementation ist die Ausrichtung des räumlichen Fokus statisch vorgegeben. Flexible, d.h. situationsabhängige Festlegungen, von räumlichen Foki sind in einer späteren Implementation vorgesehen.

---

(point werden über eine interne Repräsentationsstruktur (BSP-Baum) entsprechend ihrer Distanz geordnet. Aus operationaler Sicht werden Flächen von Objekten, die weiter entfernt sind, vor solchen gezeichnet, die näher sind. Dadurch wird erreicht, daß Überlappungen korrekt gezeichnet werden (vgl. [Bau95]). Das Ergebnis ist zwar eine korrekte Darstellung der Situation bezüglich eye point und aim point, jedoch fehlt eine explizite Repräsentation der letztendlich sichtbaren Anteile der gezeichneten Flächen. Um visuelle Salienzen berechnen zu können, ist es notwendig, die Gebiete einer Fläche zu bestimmen, die von einem Blickpunkt aus sichtbar sind.


In sequentieller Weise werden situationsabhängige Mittelwerte rekursiver und linearer Merkmale etabliert (vgl. Kapitel 4.2.3.6, 4.2.3.7). Bedingt durch die Effizienz der Verarbeitung wird in der Implementation ein rechteckiger räumlicher Fokus verwendet.88 Der räumliche Fokus definieren den Ausschnitt einer Situation, in dem visuelle Salienzen bezüglich Objekten berechnet werden. Besteht eine Intention an einem Entscheidungspunkt links abzubiegen, wird der räumliche Fokus relativ zur Frontalrichtung des primären Referenzsystems nach links verschoben.89 Insgesamt sind in MOSES drei verschiedene Fokussierungsmöglichkeiten implementiert. Der Agent kann entweder nach links, rechts oder geradeaus fokussieren (vgl. [Bau95]).

Für jeden Merkmalsyp wird, entsprechend der “feature integration theory”, eine eigenständige Rasterrepräsentation (Merkmalskarte) etabliert (vgl. Kapitel 2, 4.2.3). Bezüglich einer solchen Merkmalskarte werden in einem räumlichen Fokus, entsprechend den in Kapitel 4.2.3 vorge-
stellten Funktionen, visuelle Salienzen bezüglich rekursiver Merkmale etabliert. Bezüglich dem rekursiven Merkmal *Farbe* werden visuelle Salienzen, entsprechend der an die MacAdamschem Ellipsen angepaßten, durch Abstände von Farbvektoren entsprechend dem von MacAdam vorge- schlagenen CIE-UCS-Diagramms (uniform chromaticity scale) (vgl. [Mac37], [WS82], [JW75]).


---


92In der aktuellen Version sind diese Schwellwerte statisch. Wiederum in Bezug auf ressourcenbeschränkte Umgebungen sollen in einer Erweiterung dieses Ansatzes die Schwellwerte flexibel anpaßbar sein.

93Dies steht im Gegensatz zur "feature integration theory", jedoch ist es nicht sinnvoll davon auszugehen, daß ein kleiner Ball aufgrund seiner stark vom Mittelwert der Objekte einer Situation abweichenden Breite und Höhe besonders auffällig sein soll.
4.3.3 Prozesse zur Verarbeitung räumlichen Wissens


Im deiktischen Fall werden räumliche Relationen bezüglich eines primären oder sekundären Referenzsystems etabliert. Intrinsische Relationen zwischen Objekten basieren auf der Verwendung eines virtuellen Referenzsystems. Da für die Etablierung räumlicher Relationen Funktionen verwendet werden, die grundsätzlich von einem im Referenzobjekt verankerten Referenzsystem ausgehen (vgl. Kapitel 4.2.4.1 und [Her96], [Gap94]), wird aus Implementationsgründen bei der Etablierung der räumlichen Relationen #in-front-of# und #behind# ein extrinsisches Referenzsystem und für die Relationen #left-of# und #right-of# ein um 180° gedrehtes Referenzsystem angenommen. Eine solche Verwendungsweise von virtuellen Referenzsystemen bildet die Funktionen nach, die in Kapitel 4.2.4.1 vorgestellt worden sind.

Räumliche Relationen zwischen Objekten werden als Listen verwaltet, die aus dem Namen, den Referenzen auf die Objekte und dem Anwendbarkeitsgrad der Relation bestehen.

![Abbildung 4.73: Die Bedienungssoberfläche des Systems MOSES-SPEAK](image)

Das Subsystem MOSES-SPEAK ist sowohl in MOSES als Prozeß über einen Agenten integriert, als auch als eigenständiges System verwendbar (vgl. Abbildung 4.73). In MOSES-SPEAK erfolgt der dynamische Aufbau räumlichen Wissens, der Auswahl von Inhaltsstrukturen und deren Umsetzung in Phrasenstrukturen. Weiterhin wird der geschätzte Zeitrahmen $T_{\text{guess}}$ angezeigt. Das Verhalten von MOSES-SPEAK ist über Parameterstellungen modifizierbar (vgl. [Pau95]).

Ebenso wie die Bedienungssoberfläche von MOSES verfügt MOSES-VIEW über die Möglichkeit, eine inkrementelle Wegbeschreibung durch Aktivierung der Schaltfläche START zu starten. Daraufhin erscheint in der kleinen Karte rechts unten das Netz, in dem der Benutzer den Start- und Zielknoten auswählen kann. Auf der Oberfläche wird links oben jeweils die aktuell ausgewählte

---

94 Das Subsystem MOSES-SPEAK ist von Joachim Paul implementiert worden ([Pau95]).
4.3. IMPLEMENTATION


4.3.4 Prozesse zur Generierung natürlicher Sprache


4.3.5 Beispiel einer automatisch generierten inkrementellen Wegbeschreibung

In Abbildung 4.74 ist eine Beispielsequenz abgebildet. Für dieses Beispiel ist MOSES als Fußgänger konfiguriert worden, der über gute sprachliche Fähigkeiten verfügt. Weiterhin integriert MOSES nach Möglichkeit Landmarken in seinen Beschreibungen.


$^{95}$Gute Beschreibungen unterscheiden sich von normalen darin, daß sie grammatisch korrekt sind. Die Mehrzahl der Probanden generierten jedoch Beschreibungen, die teilweise nur unzureichend die Grammatik der deutschen Sprache berücksichtigten (vgl. das einleitende Beispiel in Abschnitt 1.1). Implementatorische Aspekte hierzu sind in [Pan96] genauer beschrieben.
Abbildung 4.74: Sequenz einer von MOSES generierten inkrementellen Wegbeschreibung

Abbildung 4.75: Sequenz einer von MOSES generierten inkrementellen Wegbeschreibung bei erhöhter Geschwindigkeit.
Kapitel 5

Erzielte Ergebnisse, Technologische Betrachtung und Erweiterungen


5.1 Wissenschaftlicher Beitrag dieser Arbeit

Eine verbale Beschreibung dessen, was via visueller Perzeption wahrgenommen wird, ist eines der komplexesten aber gleichzeitig auch interessantesten Fragestellungen aus Sicht der Künstlichen Intelligenz und anderer kognitionswissenschaftlich orientierter Forschungsrichtungen. Da dieser gesamte Problembereich momentan als viel zu komplex erscheint, um als Ganzes betrachtet werden zu können, erscheint es als einzige sinnvolle Herangehensweise, Teilprobleme zu betrachten. Eingebettet in diesen Kontext, ist es erstmals für die Domäne der inkrementellen Wegbeschreibungen durch das vorliegende formale Modell gelungen, eine Vielzahl von Leistungen zu integrieren, welche es erlauben, komplexe Raumbeschreibungen bei gleichzeitiger Bewegung des Sprechers zu generieren. Zusätzlich wurde anhand der verwendeten Methodologie exemplarisch aufgezeigt, in welcher Weise empirische Untersuchungen bei der Entwicklung formaler Modelle und letztendlich von Implementationen verwendet werden können.

Folgende Punkte fassen den wissenschaftlichen Beitrag dieser Arbeit zusammen:

1. Empirie - Modellbildung - Implementation

Schritt in der Verwendung dieser Methodologie dar. Besonders ist hierbei die Kombination aus Feld- und Laboruntersuchung hervorzuheben, die speziell an die zu untersuchenden Fragestellungen angepaßt worden ist. Weitere Arbeiten müssen zeigen, wie tragfähig diese Methodologie ist.

2. Sprachliche Referenzbildung in visuellen Daten


3. Ressourcenadaptives Verhalten: Berücksichtigung von zeitlichen Restriktionen

Ein weiterer Schwerpunkt dieser Arbeit liegt darin zu zeigen, wie die Generierung von inkrementellen Wegbeschreibungen auf Beschränkungen angepaßt wird, die durch die Umgebung bzw. der Bewegung des Agenten in der Umgebung determiniert sind. Mit dieser Arbeit ist erstmals untersucht worden, welche Auswirkungen zeitliche Beschränkungen auf die Generierung von Wegbeschreibungen haben. Zwar ist die zugrundeliegende empirische Untersuchung mit Fußgängern durchgeführt worden, es ist jedoch zu vermuten, daß die gefundenen zeitkategorialen Effekte in entsprechender Weise auch für andere Bewegungsarten gelten. Am Beispiel der Ressource Zeit ist untersucht worden, wie eine externe, d.h. durch die Umgebung vorgegebene Ressource, vom Agenten internialisert und verwendet wird. Basierend auf empirischen Untersuchungen ergaben sich fünf Zeitrahmen, zu denen die Probanden unterschiedlich strukturierte Beschreibungen gaben.

4. Erweiterung des Konzeptes des räumlichen Referenzsystems auf dynamische Aspekte


5. Intentionsgesteuerte Auswahl von Referenzobjekten mittels visueller Salienz

Eine der zentralen Fragen im Bereich der Verbindung visueller Datenverarbeitung und Sprachgenerierung ist, wie Referenzen zwischen Repräsentanten auf der Sprachebene mit Repräsentandenauf der visuellen Ebene verknüpft werden können. Im Kontext von Raumbeschreibungen liegt der Fokus der Untersuchungen vor allem auf der Auswahl von geeigneten Referenzobjekten. Dies ist in bisherigen Arbeiten meist unter der Prämisse einer


6. Inhaltskategorien inkrementeller Wegbeschreibungen

Durch Analyse des empirisch erhobenen Korpus inkrementeller Wegbeschreibungen ließen sich fünf Kategorien extrahieren, die in charakteristischer Weise zu Inhaltsstrukturen zusammengesetzt werden. Inhaltsstrukturen werden dynamisch in Abhängigkeit von der zur Verfügung stehenden Zeit verwendet. Entsprechend den damit verbundenen relativen Häufigkeiten dienen Inhaltsstrukturen als semantische Repräsentation für die Sprachgenerierung inkrementeller Wegbeschreibungen.

Zusammenfassend leistet das vorgeschlagene formale Modell einen Beitrag im Hinblick auf die Forderungen nach Adaptivität, Flexibilität, Effizienz und Effektivität am Beispiel inkrementeller Wegbeschreibungen.

5.2 Einordnung dieser Arbeit in den technologischen Kontext


Die Folge dieser Entwicklung ist die Forderung, die in gleicher Weise für Programmiersprachen, Tabellenkalkulations- und Textverarbeitungsprogramme gilt, daß Programmsysteme sich heutzutage an die funktionalen, aber auch ergonomischen Bedürfnisse des Benutzers adaptieren sollen. Die Bedürfnisse der Benutzer lassen sich in zwei Klassen einteilen. Erstens gibt es die Klasse der Bedürfnisse, die von einer hinreichend großen Menge von Personen geteilt werden. Beispielsweise zeigt es sich, daß Benutzer auf Betriebssystemebene bevorzugt mit graphischen Elementen

Da man heute weit davon entfernt ist, zu verstehen, in welcher Weise Programme einen Benutzer maximal unterstützen, ist man auf Approximationen angewiesen, die schrittweise verfeinert werden. Ein besseres Verständnis über die Bedürfnisse erhält man durch empirische Untersuchungen. Eine grundlegende Annahme ist dabei, daß ein Programmsystem um so besser ist, je geringer der Aufwand ist, den eine Person betreiben muß, um eine bestimmte Aufgabe zu lösen. Demzufolge sind Programme so zu konzipieren, daß sie dem Benutzer situationsabhängig die Information liefern, die das Lösen einer Aufgabe mit minimalem Aufwand seitens des Benutzers unterstützen. Dies erfordert ein Verständnis besonders der kognitiven Prozesse, aber auch der Präferenzen, die sich z.B. kulturbedingt herauskristallisiert haben. Ein Beispiel hierzu ist die Bevorzugung der vis-à-vis Situation zwischen Kommunikationspartnern (vgl. Kapitel 2.5).


5.2. EINORDNUNG IN DEN TECHNOLOGISCHEN KONTEXT

5.2.1 Navigationshilfssystem in Realumgebungen

Beschreibungen von Wegen werden im alltäglichen Leben häufig gegeben. Wege, die man häufig benutzt, kennt man „im Schlaf“. In fremden Umgebungen besitzt man ein derartiges räumliches Wissen jedoch nicht.


Im Bereich des Transportes und der Navigation, besonders zu militärischen Zwecken, sind Technologien entwickelt worden, die computerbasierte Lösungen von Orientierungs- und Navigationsproblemen erst ermöglichen:

1. Global positioning systems (GPS)
2. Automatische Fahrzeuglokalisierung (AVL)
3. Pfadsuchalgorithmen
4. Elektronische Karten und andere Datenmengen
5. Fahrzeugbasierte Navigations- und Fahrerinformationssysteme

In Bezug auf die ersten drei Punkte findet man eine allgemeine verfügbare und ausgetestete Technologie vor, wohingegen die letzten beiden Punkte die beiden größten Herausforderungen der aktuellen Forschung und Entwicklung darstellen. Man ist heute weit davon entfernt, über allgemein zugängliche und ausreichend genaue elektronische Karten zu verfügen. Ein damit verbundenes Problem, welches besonders Europa und vor allem Deutschland große Schwierigkeiten bereitet, ist der Umstand, daß die Daten in diesen Regionen von privatwirtschaftlichen Firmen erhoben und digitalisiert werden. Dadurch werden in Europa die Daten von mehreren
Firmen, wie EGT und TeleAtlas, parallel erhoben. Dies führt zur Etablierung verschiedener Datenstandards, was die Verwendbarkeit der Daten einschränkt. In den USA dagegen werden die Grunddaten vom “U.S. Bureau of Censor” zur Verfügung gestellt.


Das allgemeine Navigationsproblem läßt sich in acht Teilmodule unterteilen ([FHW93], [NCG93]):

1. Orientierung (Wo befindet man sich?)
2. Pfadsuche (Welcher Weg führt dort hin?)
3. Navigationshilfe (Wie kommt man dort hin?)
4. Weginformation (Was befindet sich entlang des Pfades?)
5. Pfadzustand (Wie ist der Weg?)
6. Fahrtenplanung (Wann kommt man an?)
7. Fahrtenzuteilung (Wer kommt dort als erstes an?)
8. Modusauswahl (Welche Fortbewegungsart wird gewählt?)

Jede dieser Teilmodule wird von verschiedenen Navigationssystemen in unterschiedlicher Weise verwendet. So wurden im Rahmen des IVHS-Programms 27 verschiedene Navigationsaufgaben im Kontext von Fahrradnavigationssystemen klassifiziert (vgl. [CCC+94]):

1. Information vor Beginn einer Fahrt: Lieferung von Reiseinformation für einen kompletten Bereich von multimodalen Transport- und Beförderungsoptionen auf Nachfrage vor Beginn der Fahrt
2. Fahrgebende Fahrer-Information: Aktuelle Verkehrsinfomation für den Fahrer
3. Fahrgebende Reiseinformation: Aktuelle VerkehrsInformation über multimodale Beförderungsoptionen während der Reise
4. Reiseservice Information: Information über den gesamten Bereich reiserelaterer Serviceleistungen und Einrichtungen auf Anfrage vor und während der Reise

¹Dedicated Road Infrastructure for Vehicle Safety in Europe
²Intelligent Vehicle Highway Systems
³Vehicle Information and Communication System
5. **Fahrtenleitung**: Wegbeschreibungen im statischen und dynamischen Modus entsprechend benutzerspezifischer Einstellungen.

6. **Fahrtenidentifikation und -reservierung**: Information über Mitfahrangebote

7. **Unfall- und Behinderungs-Management**: Information über Verkehrsfluß und Umgebungs- 
zustände sowie Koordination bei Unfällen

8. **Reisemanagement**:
   Enthält Fahrbahnverkehrsregelung, Parkmanagement-Systeme, Luftverschmutzung/Emissionsaufzeichnung, etc.

9. **Verkehrskontrolle**: Verkehrüberwachung und -vorsagen zur optimalen Steuerung des Verkehrsdurchflusses

10. **Elektronische Abrechnungssysteme**: Automatische Abrechnung von Gebühren von einem 
    Konto für die Benutzung einer Straße

11. **Elektronische Fahrzeuginspektion**: Automatische Inspektion von Fahrzeug- und Fahrerin- 
    formation sowie Frachtgutscheinen

12. **Elektronische Fahrzeuginspektionen am Wegesrand**: (vgl. mit verkürzten TÜV-Kontrollen)

13. **Kommerzielle Fahrzeugverwaltungsprozesse**: Elektronischer Erwerb von Transportgenehmigungen, automatische Aufzeichnung und finanztechnische Weitergabe von Kilometerstand und Verbrauch etc.

14. **Fahrzeuggestützte Sicherheitsüberwachung**: Überwachung des Zustands des Fahrzeugs, der 
    Fracht und des Fahrers während der Fahrt.

15. **Kommerzielles Flotten-Management**: Identifikation, Lokalisation von Fahrzeugen, Fahren- 
    leitung, Flottenverwaltung und Echtzeitstatus zur Unterstützung der Flottenoperationen

16. **Öffentliches Verkehrsmittelmanagement**: Verwaltung von intermodalen, öffentlichen Ver- 
    kehrsmitteln in Echtzeit, Serviceplanung, etc.

17. **Personalisierter öffentlicher Verkehr**: Auswertung und Nutzung von Kundenanfragen, Ko- 
    ordination von Fahrtenanfragen etc.

18. **Notfallbenachrichtigung und persönliche Sicherheit**: Benachrichtigung von verantwortlichen 
    Stellen, Management von Einsatzplänen etc.

19. **Sicherheit öffentlicher Verkehrsmittel**: Überwachung öffentlicher Verkehrsmittel und Ein- 
    richtungen durch visuelle und auditive Sensoren

20. **Management von Einsatzfahrzeugen**

21. **Longitudinale Kollisionsvermeidung**: Auffahrunfallwarnung- und -kontrolle, autonome bzw. 
    kooperative und intelligente Fahrzeugsteuerung, Überholwarnung, etc.

22. **Laterale Kollisionsvermeidung**: Fahrbahnwechsel/Blinder Punkt Kollisionsvermeidung, 
    Fahrbahnverfolgungskontrolle

23. **Kreuzungskollisionswarnung und -kontrolle**

24. **Kollisionsvermeidung durch Verbesserung der visuellen Sicht**: Aktive Verbesserung der 
    visuellen Sicht zur Vermeidung von Kollisionen bei schlechten Sichtverhältnissen
25. Übertretungsalarms: Aktive Fahrzeug- und Straßenzustandsüberwachung zur Alarmierung des Fahrers

26. Anwendung von Sicherheitsmaßnahmen unmittelbar vor einer Kollision


Diese Anwendungen sind nicht strikt voneinander getrennt, sondern bauen teilweise aufeinander auf. Einige hören sich heutzutage noch sehr futuristisch an und überschreiten mit Sicherheit deutsche Datenschutzgesetze. Sie zeigen jedoch das allgemeine Potential von Navigationssystemen an und geben einen Hinweis darauf, was in Zukunft zumindest auf dem amerikanischen Markt angeboten werden wird. Beispielsweise sind testweise bereits zwei elektronische Abrechnungssysteme in den USA im Einsatz.

Jede einzelne Anwendung basiert auf den vorher indizierten Anforderungen in unterschiedlicher Weise (vgl. [NCG93], [CCC+94]). Vor diesem Hintergrund der Komplexität der Aufgaben, die im Kontext intelligenter Verkehrssysteme auftreten, läßt sich ein inkrementelles Wegbeschreibungssystem, so wie es durch MOSES vorgestellt worden ist, als System zur Fahrtenleitung verstehen.

Eine natürliche technische Verwendung inkrementeller Wegbeschreibungen erfolgt im Zusammenhang mit intelligenten Fahrernavigationssystemen ([MWH93]). Von verschiedenen Firmen werden erste Fahrernavigationssysteme angeboten, die den Namen wirklich verdienen. So bieten BLAUPUNKT/BOSCH (TravelPilot) und PHILIPS (Carin) zwei sehr ähnliche Systeme an, die dem Fahrer über einen kleinen Monitor auf der Konsole des Wagens die lokale Konfiguration der Straßenabschnitte anzeigt und Richtungsangaben mittels Pfadsymbolen indiziert, sowie verbal eine kurze Beschreibung der nächsten Aktion am nächsten Verzweigungspunkt angibt (vgl. Abbildung ??). Im System von BLAUPUNKT/BOSCH wird die Position des Wagens durch Abgleich von räumlichen Daten vorgenommen, die via GPS empfangen, von einer CD abgerufen, sowie durch Radsensoren und per Kompaß ermittelt werden. Dies ist notwendig, da die Daten des GPS vom amerikanischen Verteidigungsministerium aus militärischen Gründen gestört werden, was zu einer Herabsetzung der Genauigkeit der Koordinaten führt. Dadurch erlauben die empfangenen Daten nur noch die Bestimmung der Lokation bis auf ca. 10 Meter genau.


Generell ist bei der Entwicklung von Navigationsschnittstellen darauf zu achten, daß der Fahrer nur relevante Information zur richtigen Zeit erhält ([FRH92]). Bedingt durch identische Prämisse kann MOSES in eingeschränkter Weise als Fahrernavigationssystem verstanden werden. MOSES erweitert momentan auf dem Markt befindliche Systeme dadurch, daß es über eine dynamische, inkrementelle Generierung von Beschreibungen verfügt und diese in Abhängigkeit
von der Situation präsentiert. Im Gegensatz zu Navigationssystemen, die sich momentan im Einsatz befinden, ist MOSES in der Lage Landmarken aus einer Situation in die Beschreibung zu integrieren. Zwar wird aktuell das Kartenmaterial um Landmarkeninformation erweitert, jedoch löst dies nicht das Problem, ob eine Landmarke überhaupt in der jeweiligen Situation sichtbar ist. Um Landmarken in solchen Navigationssystemen verwenden zu können ist es entweder notwendig, die Information über die Sichtbarkeit von Landmarken mit in der Karte zu speichern oder aber, wie in MOSES , dies über die Auswertung der visuellen Information zu lösen. Die Effizienz, die durch die erste Variante erreicht wird, geht zulasten der Flexibilität und Gültigkeit der Beschreibung. Möglicherweise kann eine Kombination aus beiden Ansätzen einen geeigneten Lösungsansatz bieten.


In ähnlicher Weise lassen sich derartige Navigationssysteme als Orientierungshilfe für komplexe Umgebungen konstruieren, wie z.B. für Flughäfen, Bahnhöfe, Museen und Krankenhäuser. Ein weiteres Gebiet eröffnet sich im Hinblick auf Rehabilitationsmaßnahmen von gehirnneurologisch geschädigten Personen, die nur noch über eingeschränkte räumliche Orientierungsfähigkeiten in ihren alltäglichen Umgebungen besitzen. Der Einsatz von Navigationssystemen ermöglicht es Patienten mit bestimmten Schädigungen, ihre Umgebung teilweise wieder zu erlernen und die Abhängigkeit von Betreuern zu verringern, was nicht zuletzt ein finanzielles und personelles Problem ist ([HM94]).


Neben dem Einsatz von Navigationssystemen bezüglich realen Umgebungen besitzen sie auch Anwendungspotential in virtuellen Umgebungen. Dieses soll im weiteren kurz diskutiert werden.

5.2.2 Navigationshilfe in virtuellen Umgebungen

In virtuellen Umgebungen werden räum-zeitliche Metaphern verwendet, um dem Benutzer einen möglichst natürlichen Zugang zu den Daten bzw. zur Funktionalität des Systems zu erlauben. Beispielsweise ist der Mülleimer der MacOS-Oberfläche mehr als nur ein Symbol. Um eine Datei zu löschen, „nimmt“ der Benutzer sie auf, „geht“ zum Mülleimer und „wirft“ die Datei

Für einen Überblick entsprechender Forschungsarbeiten zum Thema Neglect sei auf die Sonderausgabe der Zeitschrift Neurophysiological Rehabilitation verwiesen ([HM94]).


Im Hinblick auf die kommerzielle Nutzung solcher Technologien zeigt sich, daß herkömmliche

---

5Bezogen auf das WWW sind Lokationen am echten mit WWW-Seiten zu unifizieren.
6Die Verwaltung von WWW-Seiten in strukturierten Listen, wie in Netscapes Navigator 2.0, ist hierzu nur ein Anfang.

Vor diesem Hintergrund sind Wegbeschreibungen im erweiterten Sinne von hervorragendem Interesse. Alle Elemente, aus denen sich Wegbeschreibungen konstituieren, finden sich i.a. in ausreichend komplexen, virtuellen Welten wieder. Die metaphorische Verwendung dieser Elemente ist zwar in gewisser Weise an die jeweilige Umgebung anzupassen, verliert jedoch nicht die grundlegende Bedeutung. Durch Anpassung der Elemente läßt sich MOSES als Navigationsystem in einem virtuellen Kaufhaus oder dem WWW allgemein verwenden. Information über Landmarken kann dazu verwendet werden, dem Benutzer die aktuelle virtuelle Lokation zu beschreiben. Verzweigungspunkte können einerseits im obig beschriebenen Sinne eine Reise durch ein Kaufhaus beschreiben. Andererseits kann dem Benutzer auch kurz beschrieben werden, was sich hinter den Verzweigungen befindet.

5.3 Erweiterungsmöglichkeiten

Abschließend soll aufgezeigt werden, in welcher Weise das vorgeschlagene Modell unmittelbar, aber auch mittelbar erweitert werden kann.

1. Hierarchische Beschreibungen:


2. Pfadrelationen

Wie in Kapitel 2.3.2 eingehend erwähnt, wird zwischen topologischen, projektiven und
pfadbezogenen, räumlichen Relationen unterschieden. In dem Modell wurden die ersten beiden Relationenarten verwendet, um eine räumliche Vorstellung über eine Umgebung aufzubauen. Pfadrelationen erweitern diese Vorstellung dahingehend, daß ein Objekt durch eine Trajektorie abstrahiert wird, der eine räumliche Bedeutung zugeordnet wird. Bei Abbiegenvorgängen wird an Verzweigungspunkten den Objekten keine Trajektorie, sondern eine statische Punkt-zu-Punkt-Beziehung zugeordnet, wie sie durch topologische und projektierte räumliche Relationen repräsentiert werden. Pfadrelationen werden auf sprachlicher Ebene vor allem für unterstützende Beschreibungen während des Zeitrahmens $t_1$ sprachlich indiziert, wie zum Beispiel in der folgenden Beschreibung: „Gehen Sie an der Kirche entlang.“

3. Bewegungsverben


4. Kognitive Karten


5. Sprachgesteuerte Aufmerksamkeitsfokussierung


6. Vollständiger Sprecher-Hörer-Dialog und Dialoggedächtnis

Bedingt durch die Fokussierung von Beschreibungen an Verzweigungspunkten, ist der Aspekt der Kommunikation in dieser Arbeit in den Hintergrund getreten. Wegbeschreibungen sind allgemein jedoch dialogisch. Damit ist gemeint, daß der Rezipient nachfragt, um eine dediziertere Beschreibung zu erhalten oder aber um sich zu versichern, daß er noch auf dem richtigen Weg ist. In solchen dialogischen Kontexten muß der Sprecher über

\footnote{Einen möglichen Ansatz zur Modellierung von Bewegungsverben findet sich in [Ma90].}
ein Dialoggedächtnis verfügen, um einerseits die Aussagen des Rezipienten zu verstehen und um andererseits auf Vorhergesagtes referenzieren zu können. Weiterhin ist ein geeignetes Benutzermodell zu fordern, welches es dem Sprecher erlaubt, in geeigneter Weise auf die (Nach-) Fragen des Rezipienten zu reagieren. Dies bedingt eine Hierarchisierung des räumlichen Wissens und geeigneter Zugriffsprozessen. Beispielsweise ist der Weg für einen Einheimischen anders zu beschreiben, als für einen Auswärtigen.

7. Emotionen


8. Anwendung auf andere Domänen


Abschließend möchte ich diese Arbeit mit einem Zitat von Steven Vere und Timothy Bickmore (vgl. [VB90]).

"At one time it was thought that intelligence and AI had something to do with generality. The agent is subject to criticism from researchers specializing in one intelligence component, who may be able to point out that their one component works better than the corresponding component of an integrated agent. However, agents deserve credit for their "horizontal depth". Like an athlete in the decathlon, an agent should be able to earn points for its performance in each cognitive event, without having to beat all of the narrower systems in their one speciality event... Researches specializing in a particular component of intelligence should realize that an integrated agent is in fact a validation and justification of their own work."

---

8Diesen interessanten Hinweis verdanke ich Xavier Briffault.
Literaturverzeichnis


In *Speech and Natural Language Workshop*. DARPA, 1989.

[FO91] P. F. Fisher and T. M. Orf. An investigation of the meaning of near and close on a


[Fra91] A. U. Frank. Qualitative spatial reasoning with cardinal directions. In H. Kaindl,

[Fre86a] G. Frege. *Begriffsschrift*, eine der arithmetischen nachgebildete formelsprache des
reinen denkens (gekürzter nachdruck). In K. Berka and L. Kreiser, editors, *Logik-
Texte: Kommentierte Auswahl zur Geschichte der Modernen Logik (vierte Auflage)*,

Kommentierte Auswahl zur Geschichte der Modernen Logik (vierte Auflage)*, pages

[FRH92] W. Fastenmeier, G. Reichart, and R. Haller. Welche information brauchen fahrer
wirklich? In *Das Mensch-Maschine-System im Verkehr*, pages 15–33. VDI Verlag,

Boff, L. Kaufman, and J. P. Thomas, editors, *Handbook of Perception and Human

[FS92] W. Finkler and A. Schauder. Effects of incremental output on incremental natural

[FTC92] N. Franklin, B. Tversky, and V. Coon. Switching points of view in spatial mental


and W. Kuhn, editors, *Spatial Information Theory*, Lecture Notes in Computer
LITERTURVERZEICHNIS


LITERATURVERZEICHNIS


