

Wie komme ich zum Finanzamt? Zur automatischen Generierung sprachlicher Beschreibungen aus visuellen Daten

Wolfgang Wahlster
Fachbereich Informatik, Universität des Saarlandes

1. Einleitung

Es gehört sicherlich zu den schwierigsten Informationsverarbeitungsprozessen, die Menschen alltäglich durchführen, einen visuell wahrgenommenen Vorgang sprachlich so zu beschreiben, daß sich ein Gesprächspartner, der den Vorgang nicht beobachten konnte, ein Bild von dem Geschehen machen kann.

Ein Teilziel der Forschungen im Bereich der Künstlichen Intelligenz (KI) ist es, Programmsysteme zu entwickeln, die visuelle Information sprachlich umsetzen können. Das Projekt VITRA, das wir seit 1985 im Sonderforschungsbereich 314 'Künstliche Intelligenz - Wissensbasierte Systeme' verfolgen, soll einen Beitrag zur Grundlagenforschung auf dem Gebiet der Kopplung bildverstehender und sprachgenerierender Systeme leisten.

Mit VITRA wird eine formale Rekonstruktion wichtiger Aspekte des Zusammenspiels zwischen 'Sprechen und Sehen' angestrebt, die eine Voraussetzung für intelligente Computersysteme mit 'Augen und Ohren' darstellt.

Langfristig verfolgt man mit dieser Forschungsrichtung zwei Hauptziele:

- (a) Die komplexen Informationsverarbeitungsprozesse des Menschen, die der Interaktion von Sprachproduktion und visueller Wahrnehmung zugrundeliegen, sollen mit informatischen Mitteln exakt beschrieben und erklärt werden.
- (b) Durch die sprachliche Bildbeschreibung sollen dem Benutzer die Ergebnisse eines bildverstehenden Systems besser zugänglich und verständlich gemacht werden.

Charakteristisch für die KI-Forschung ist, daß neben dem kognitionswissenschaftlichen Erkenntnisinteresse (a) auch eine ingenieurwissenschaftliche Zielsetzung (b) verfolgt wird.

Ein großer praktischer Vorteil der sprachlichen Bildbeschreibung besteht in der Möglichkeit zur anwendungsspezifischen Wahl unterschiedlicher Verdichtungsgrade für visuelle Information (vgl. Wahlster 1989). So kann die beispielsweise in der Medizintechnik, der Fernerkundung und der Verkehrssteuerung anfallende Flut von Bilddaten nur noch maschinell bewältigt werden. Im Gegensatz zu einer Repräsentation der Verarbeitungsergebnisse der digitisierten Bildfolgen in Form von Computergraphiken kann eine sprachliche Bildbeschreibung dem Anwender mehr Information in weniger Zeit liefern. Wenn ein KI-System in der Lage ist, das Interpretationsergebnis für eine Bildfolge in einer medizinischen Anwendung mit "Verengung der linken Nierenarterie" zu beschreiben, so kann der Arzt diese Aussage zunächst direkt in den diagnostischen Zusammenhang einordnen und kann dann später bei Bedarf gezielt auf Ausschnitte relevanter Einzelbilder zurückgreifen.

In Umkehrung eines alten chinesischen Sprichwortes gilt in solchen Situationen 'Ein Wort sagt mehr als 1000 Bilder'.

Derzeit ist man von einem universell einsetzbaren KI-System, das beliebige Bildfolgen sprachlich beschreibt, noch sehr weit entfernt und muß sich bei der Systementwicklung jeweils auf eingeschränkte Diskursbereiche konzentrieren. Im Projekt VITRA werden vier verschiedene Diskursbereiche und zwei unterschiedliche Kommunikationssituationen betrachtet, um möglichst frühzeitig die Übertragbarkeit der entwickelten Konzepte und Methoden auf andere Domänen prüfen zu können:

Kommunikationssituation K1: Beantwortung natürlichsprachlicher Anfragen über räumliche Relationen und Bewegungsverläufe nach Ablauf einer Bildsequenz sowie Wegbeschreibungen

Diskursbereich D1: CITYTOUR - Stadtplanausschnitt von Saarbrücken mit Trajektorien bewegter Objekte

Diskursbereich D2: UNITOUR - Lageplan des Campus der Universität des Saarlandes

Diskursbereich D3: DURLACHER TOR - Straßenverkehrsszene aus Karlsruhe mit Trajektorien bewegter Objekte

Kommunikationssituation K2: Simultane Berichterstattung über beobachtete Ereignisse während des Ablaufs einer Bildsequenz

Diskursbereich D4: SOCCER - Ausschnitte aus Fernsehübertragungen von Fußballspielen (vgl. Herzog et al. 1989)

Während in (K1) die Rolle des KI-Systems der eines Ortskundigen ähnelt, der Auskünfte erteilt, ist sie in (K2) mit einem Radioreporter vergleichbar. Für beide Typen von Situationen gibt es zahlreiche realistische Anwendungsszenarios. So könnte beispielsweise ein Biologe im ersten Fall aufgrund einer Folge ausgewerteter Luftbilder fragen "Wo wurden Schädigungen von Birkenbeständen festgestellt?". Im zweiten Fall erwartet z.B. der Bediener eines Leitstandes für ein komplexen technisches System eine Beschreibung einer sich anbahnenden Fehlfunktion oder eine Warnung vor einer potentiellen Betriebsstörung.

In diesem kurzen Beitrag soll die prinzipielle Arbeitsweise unseres Systems am Beispiel der Generierung von Wegauskünften innerhalb der Saarbrücker Innenstadt gezeigt werden. Dabei werden viele Vereinfachungen vorgenommen und es kann nur ein sehr kleiner Ausschnitt der Funktionalität des Gesamtsystems vorgestellt werden.

2. Probleme der automatischen Wegauskunft

Eine für den Menschen alltägliche Aufgabe besteht darin, Ortsunkundigen zu beschreiben, auf welchem Weg sie zu einem bestimmten Ziel gelangen. Es ist selbstverständlich, daß eine automatische Wegbeschreibung schon alleine wegen der extremen Kombinatorik möglicher Anfragen nicht aus der Suche und Präsentation einer endlichen Zahl vorgefertigter Auskünfte bestehen kann, sondern ein generatives Verfahren realisiert werden muß, das potentiell unendliche viele Wegauskünfte synthetisieren kann.

Es ist nicht ratsam, sich bei der Konstruktion eines KI-Systems zur Generierung von Wegbeschreibungen vollständig am Menschen zu orientieren, da mehrere sozialwissenschaftliche Studien zeigten, daß nur ca. 50% der spontan und ohne Verwendung einer Karte erzeugten Wegauskünfte von Menschen brauchbar sind. Dies ist u.a. darauf zurückzuführen, daß die sog. kognitiven Karten des Menschen als Grundlage der Wegbeschreibung sehr unpräzise und teilweise sogar objektiv falsch sind. Mit dem Begriff der kognitiven Karte bezeichnet man die Repräsentation räumlicher Information im menschlichen Gedächtnis. Da schon allein wegen der Größe nicht die gesamte Umwelt durch unmittelbare visuelle Wahrnehmung erfaßt werden kann, kombiniert der Mensch viele direkte Wahrnehmungen zu einer notwendig unpräzisen und partiellen räumlichen Gesamtvorstellungen in kognitiven Karten.

Eine Vielzahl psychologischer Untersuchungen hat gezeigt, daß kognitive Karten erhebliche Verzerrungen enthalten, Richtungsänderungen meist als rechte Winkel angenähert werden, Abstände zwischen Richtungswechseln oft als äquidistant angenommen werden, Straßen begradigt und nach Haupthimmelsrichtungen ausgerichtet werden.

Andererseits ist eine rein mathematische Lösung als Suche nach einem kürzesten Weg auch keine für den Systembenutzer akzeptable Lösung. Jeder Informatik-Student kennt den Dijkstra-Algorithmus, der in einem Wegegraphen den kürzesten Pfad zwischen zwei Knoten des Graphen findet. Es ist bekannt, daß dieses Verfahren bei einer geschickten Speicherung des Graphen im Rechner im ungünstigsten Fall größenordnungsmäßig $n \log(n)$ Berechnungsschritte braucht, wobei n die Zahl der Knoten im

Graphen ist. Wesentliche Voraussetzung für eine erfolgreiche Wegbeschreibung ist jedoch nicht nur ihre Eindeutigkeit und Verifizierbarkeit, sondern auch, ob die Merkfähigkeit des Adressaten mitberücksichtigt wurde (vgl. Hoepfner et al. 1989). Es ist ein Charakteristikum guter Wegebeschreibungen, daß sie häufig einen etwas längeren Weg beschreiben, der aber einfacher zu beschreiben, zu merken oder zu finden ist.

Hier zeigt sich ein Paradox, mit dem sich der KI-Forscher auseinandersetzt: Rechner speichern Information zuverlässig und dauerhaft, so daß Vergessensprozesse eigentlich nicht zu interessieren brauchen. Will man aber eine dem menschlichen Benutzer angepaßte und für ihn brauchbare Ausgabe erreichen, so muß man versuchen, dessen mögliches Vergessen als Performanzfehler zu antizipieren und daher letztlich doch eine - wenn auch sehr rudimentäre - Theorie des Vergessens implementieren. Das System sollte also statt 'Rechts, Rechts, Rechts, Links, Links, und dann Rechts abbiegen' besser 'Immer Rechts bis zur Tankstelle, dann zwei mal Links und an der Kirche rechts abbiegen' generieren. Bei der verständlicheren Variante wurden u.a. "Tankstelle" und "Kirche" als sog. Landmarken verwendet. Landmarken sind markante Bezugspunkte im Raum, die durch ihre Auffälligkeit oder besondere Funktion wichtige organisatorische Elemente von kognitiven Karten darstellen. Die Wegbeschreibung soll beim Auskunftsuchenden eine Raumvorstellung erzeugen, die er bei der eigentlichen Wegfindung leicht überprüfen kann. Hier zeigt sich, daß mit der sog. Benutzermodellierung ein weiteres Teilgebiet der KI von großer Bedeutung für eine automatische Wegbeschreibung ist. Denn das System muß bei der Auswahl und Verbalisierung von Landmarken das individuelle Vorwissen des Fragenden berücksichtigen. So hätte es z.B. wenig Sinn, in einer Wegbeschreibung innerhalb Saarbrückens einen Hinweis wie 'am PEKA vorbei' zu verwenden, wenn bekannt ist, daß der Fragende nicht aus Saarbrücken kommt und daher wohl kaum wissen kann, daß dies die ältere, aber für einen Saarbrücker immer noch populärere Bezeichnung für das Kaufhof-Gebäude darstellt.

3. CITYGUIDE: das Wegauskunftssystem in VITRA

CITYGUIDE (vgl. Müller 89) ist dasjenige Subsystem von VITRA, das aus beliebigen digitisierten Stadtplänen die für eine Wegauskunft relevante Information extrahiert und wahlweise in deutscher oder englischer Sprache Wegbeschreibungen generiert.

Abb. 1 zeigt einen Bildschirmabzug des hochauflösenden Graphik-Monitors der LISP-Maschine, auf der VITRA implementiert ist. Das große rechte Graphikfenster zeigt einen Ausschnitt aus dem Stadtplan von Saarbrücken, in dem CITYGUIDE den aufgrund der Benutzeranfrage gesuchten Weg markiert hat.

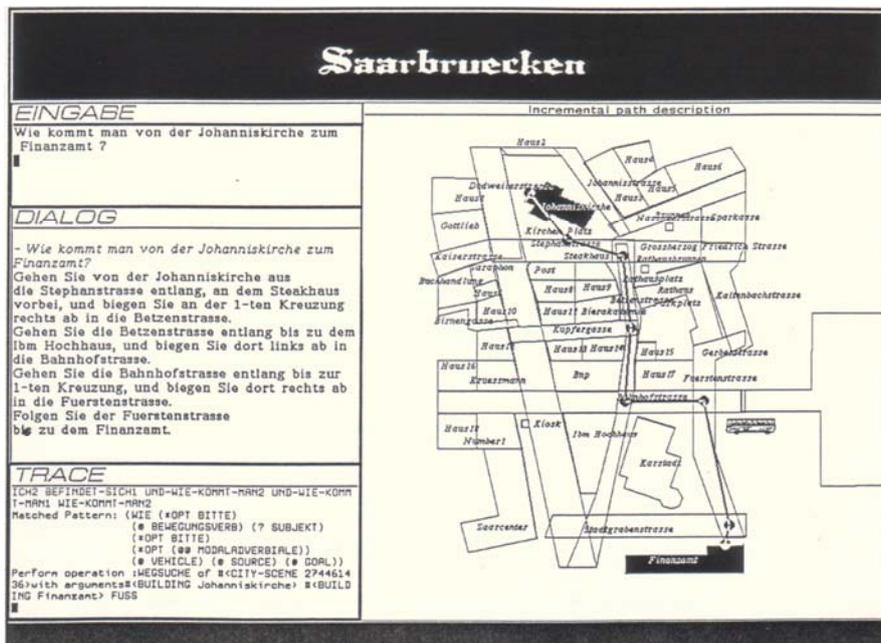


Abbildung 1

Das linke Drittel des Bildschirms enthält drei Textfenster: Eingabe, Dialog und Trace. Die in natürlicher Sprache formulierten Anfragen des Benutzers werden in das Eingabefenster eingetippt. Im Dialogfenster erscheint die Antwort von VITRA. Über einen angeschlossenen Synthesemodul für gesprochene Sprache kann die Wegauskunft auch lautsprachlich realisiert werden. Im Trace-Fenster hat der Benutzer u.a. die Möglichkeit, die internen Verarbeitungsabläufe des Systems zu verfolgen oder sich die von VITRA verwendeten Wissensquellen anzeigen zu lassen.

CITYGUIDE verwendet zur Wegsuche eine Erweiterung des Algorithmus von Dijkstra zur Suche kürzester Wege in Graphen. Deshalb muß zunächst ein entsprechender Suchgraph aus dem Stadtplan aufgebaut werden. Er enthält zusätzlich Markierungen, wer die entsprechenden Straßenabschnitte benutzen kann (aus der Menge: Fußgänger, Fußgänger mit Kinderwagen, Rollstuhlfahrer, Radfahrer, Kraftfahrzeug), so daß ein dem in der Benutzerfrage spezifizierten Fortbewegungsmittel angepaßter Weg gefunden werden kann (vgl. Abb. 2).

Die Knoten des Graphen repräsentieren folgende Stellen in der Szene:

- Kreuzungen einer Straße mit einer oder mehreren anderen: Knoten dieser Art erhalten als Namen die Liste aus den Namen der sich kreuzenden Straßen, als Koordinaten die des Mittelpunkts des Kreuzungsbereichs.
- Kreuzungen einer Straße und eines Platzes: Der Name eines solchen Knotens ist die Liste aus dem Namen der Straße und dem Namen des Platzes, die Koordinaten sind die des Mittelpunkts der Schnittfläche.
- Plätze: Ein Platzknoten ist nach dem Platz benannt und hat als Koordinaten dessen Schwerpunktkoordinaten.
- Platzschnitt- oder Berührungspunkte: Ein Knoten dieser Art erhält als Namen die Liste aus den Namen der beiden Plätze, seine Koordinaten sind die des Mittelpunkts der gemeinsamen Fläche.
- Straßenanfangs- bzw. Endpunkte, die nicht auf einer Kreuzung liegen: Ein Anfangs- oder Endknoten ist nach der Straße benannt, deren Ende er repräsentiert, seine Koordinaten sind die des Mittelpunkts zwischenden Endpunkten des linken und des rechten Randes der Straße.

Dazwischen sind nach folgenden Regeln Kanten eingetragen:

- Ein Platzknoten ist durch Kanten mit allen Knoten verbunden, in denen der Platz eine Straße schneidet.
- Einen Knoten, der eine Einmündung einer Straße in einen Platz darstellt, verbinden Kanten mit allen anderen Einmündungen in diesen Platz sowie mit dem Platzmittelpunkt und dem nächsten Knoten der Straße.
- Vom Anfangsknoten einer Straße aus gibt es eine Kante zum nächsten Knoten der Straße.
- Von einem Kreuzungsknoten gibt es Kanten zu allen benachbarten Knoten.

Dabei ist zu beachten, daß Kanten, die gemäß den möglichen Fahrtrichtungen einer Straße verlaufen, allen Verkehrsteilnehmern zur Verfügung stehen, die sich auf dem repräsentierten Wegabschnitt bewegen dürfen. Kanten entgegen der Fahrtrichtung einer Straße sind dagegen nur für Fußgänger, Kinderwagen und Rollstuhlfahrer zugelassen. Wenn eine Einbahnstraße nur von Kraftfahrzeugen benutzt werden darf, gibt es keine Kanten entgegen der erlaubten Fahrtrichtung (vgl. Abb. 2).

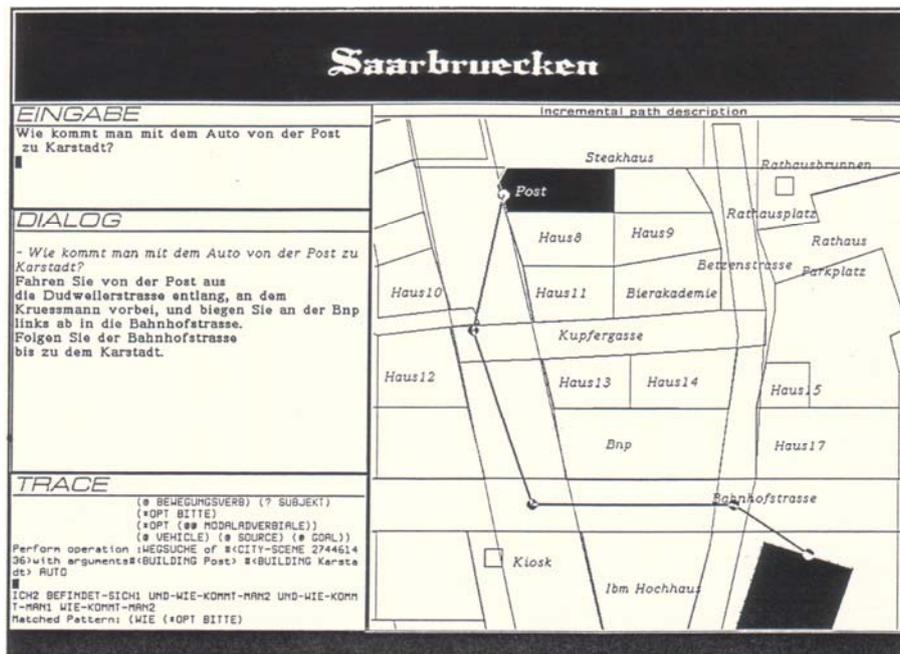


Abbildung 2

Der Graph wird aus den als Polygonen repräsentierten Objekten schrittweise konstruiert. Zuerst werden die Straßen untersucht, dann die Plätze und zuletzt die Gebäude. Die Straßen werden zunächst paarweise betrachtet. Überall dort, wo sich zwei Straßen schneiden, legt der Algorithmus einen Knoten an. Es gibt jedoch Kreuzungen, an denen mehr als zwei Straßen beteiligt sind. Um sie herauszufinden, überprüft der Algorithmus die erzeugten Knoten. Wenn er feststellt, daß zwei Knoten nicht zwei verschiedene Kreuzungen bilden, sondern eine einzige, faßt er sie zu einem Knoten zusammen.

Damit zwei Knoten eine einzige Kreuzung bilden, müssen folgende zwei Bedingungen erfüllt sein:

- Mindestens ein Straßennamen muß in den Namen beider Knoten vorkommen; sonst kann es sich nicht um eine gemeinsame Kreuzung handeln.
- Die beiden Knoten müssen bzgl. des Stadplans dicht beieinander liegen. Kriterium ist die Breite der breitesten an der Kreuzung beteiligten Straße.

Innerhalb des erzeugten Graphen wird eine Modifikation des Dijkstra-Algorithmus verwendet, der Rücksicht auf die Bewegungsmöglichkeiten des Benutzers nimmt: Es wäre z.B. unsinnig, für Rollstuhlfahrer einen Weg mit einer Treppe zu wählen. Die Kantenmarkierungen im Graph werden zu diesem Zweck getestet.

Die Beschreibungskomponente von CITYGUIDE wird von der Suchkomponente aufgerufen; dabei werden der gefundene Weg, ein Verb der Bewegung und das vom Benutzer erwähnte Fahrzeug als Parameter übergeben. Die Beschreibungskomponente generiert eine natürlichsprachliche Beschreibung des Weges unter Verwendung des angegebenen Verbs. Der in Form einer Knotenliste vorliegende Weg kann i.a. jedoch nicht als Ganzes beschrieben werden, es sei denn, er ist sehr kurz. Deshalb zerlegt ihn die Beschreibungskomponente zunächst in kleinere Teile und beschreibt diese. Die Beschreibung des gesamten Weges ergibt sich durch Aneinanderreihen der Beschreibungen der einzelnen Teile.

Die Teile sind (a) der Ausgangspunkt, (b) das Mittelstück und (c) das Ziel, von denen das Mittelstück weiter unterteilt ist, und zwar in Abschnitte, die mit jeweils einer Präpositionalphrase beschrieben werden können. Die aufgeteilte und mit weiteren Angaben versehene Knotenliste heißt Route.

Die Beschreibungskomponente gewinnt die Route aus den Knoten des Weges, indem sie die jeweils auf einer Straße oder einem Platz liegenden Knoten zu einer Liste zusammenfaßt. Die einzelnen

Wegabschnitte werden ergänzt um Landmarken, an denen sie vorbeiführen.

```
((:START Finanzamt)
 (:MITTE
  ((:STRASSE Fürstenstraße)
   (:KNOTEN (Knoten (Fürstenstraße Stadtgrabenstraße)
                Knoten (Fürstenstraße Bahnhofstraße))
    (:MERKMALE (Finanzamt))
  ((:STRASSE Bahnhofstraße)
   (:KNOTEN (Knoten (Fürstenstraße Bahnhofstraße)
                Knoten (Gerberstraße Bahnhofstraße))
    (:MERKMALE ( ) ))
  ((:STRASSE Gerberstraße)
   (:KNOTEN (Knoten (Gerberstraße Bahnhofstraße)
                Knoten (Parkplatz Gerberstraße)
                Knoten (Gerberstraße Kaltenbachstraße))
    (:MERKMALE (Haus17 Rathaus)))
  ((:STRASSE Kaltenbachstraße)
   (:KNOTEN (Knoten (Gerberstraße Kaltenbachstraße)
                Knoten (Rathausplatz Kaltenbachstraße))
    (:MERKMALE ( ) ))
 (:ZIEL Rathaus))
```

Beispiel für die interne Beschreibung einer gefundenen Route

Zuerst erzeugt die Beschreibungskomponente die Beschreibung des Ausgangsortes des Weges, dann die für den ersten Wegabschnitt. Diese beiden bilden einen einzigen Satz. Anschließend werden die Texte zur Beschreibung der weiteren Wegabschnitte generiert.

Zuletzt wird der Text zur Beschreibung des Ziels generiert. Die Methode dafür beachtet, welcher Art das Ziel ist, ob es ein Gebäude, eine Straße oder ein Platz ist, und ob der letzte Wegabschnitt eine Straße oder ein Platz war. Für die einzelnen Kombinationen erzeugt sie verschiedene Beschreibungstexte, die an die Beschreibung des letzten Wegabschnitts angefügt werden.

Die Methode, die die Beschreibung eines Wegabschnitts generiert, beschreibt auch den Übergang vom beschriebenen Wegabschnitt zum nächsten, und zwar mit folgenden Mitteln:

- Merkmal (Gebäude) an der Abbiegestelle, falls vorhanden; sonst Angabe, an der wievielten Kreuzung seit Betreten der zu verlassenden Straße abgelenkt werden soll,
- Richtung, in die abgelenkt werden muß,
- Straße oder Platz, in die bzw. den eingelenkt wird.

Die Richtung, in die abgelenkt werden muß, berechnet sich wie folgt: Es wird ein Strahl durch die beiden letzten Knoten des Routenstücks vor der Richtungsänderung gelegt und einer durch die beiden ersten Knoten des Routenstücks nach der Richtungsänderung. Der Winkel zwischen den beiden Strahlen wird berechnet, und abhängig von der Größe des Winkels erhält der Benutzer die Angabe, rechts, halb rechts, links oder halb links abzubiegen bzw. geradeaus zu gehen.

Um räumliche Beziehungen sprachlich ausdrücken zu können, verfügt VITRA über geometrische Definitionen der relevanten räumlichen Präpositionen wie "entlang", "vor", "zwischen" usw. Man kann räumliche Relationen dadurch definieren, daß man Bedingungen über räumliche Gegebenheiten einer Objektkonfiguration spezifiziert, wie z.B. Abstand zwischen Objekten, relative Lage bezüglich einer Orientierung; d.h. man kennzeichnet eine bestimmte Klasse von Objektkonfigurationen (vgl. Retzschmidt 1988). Räumliche Beziehungen zwischen Objekten lassen sich propositional durch Relationentupel folgender Form repräsentieren:

(Rel-Name Subjekt Bezugsobjekt_1 ... Bezugsobjekt_n Orientierung)

Das erste Argument bezeichnet die entsprechende räumliche Relation. Das als Subjekt bezeichnete Argument steht für dasjenige Objekt, das relativ zu einem oder mehreren Objekten, den Bezugsobjekten, (und evtl. bezüglich einer Orientierung) lokalisiert werden soll. Wir sprechen von der Anwendbarkeit eines Relationentupels, falls es sich zur Charakterisierung einer Objektkonstellation eignet. Zur Bestimmung der Anwendbarkeit erweist sich die Verwendung von Anwendbarkeitsräumen als hilfreich. Der grundlegende Gedanke dabei ist, daß man jedem Relationentupel einen Anwendbarkeitsraum zuordnet und dann prüft, in welcher mengentheoretischen Beziehung (Inklusion, Exklusion oder Überlappung) sich der vom Subjekt eingenommene Raum zu dem durch die restlichen Argumente des Relationentupels bestimmten Anwendbarkeitsraums befindet. Die Bestimmung der Anwendbarkeitsräume ist im allgemeinen nicht trivial, weil u.a. Orientierung, Ausdehnung und Form der Bezugsobjekte zu berücksichtigen sind. Weitere Schwierigkeiten ergeben sich aufgrund benachbarter Objekte, die, anschaulich gesprochen, zu einer Deformierung des Anwendbarkeitsraumes führen können.

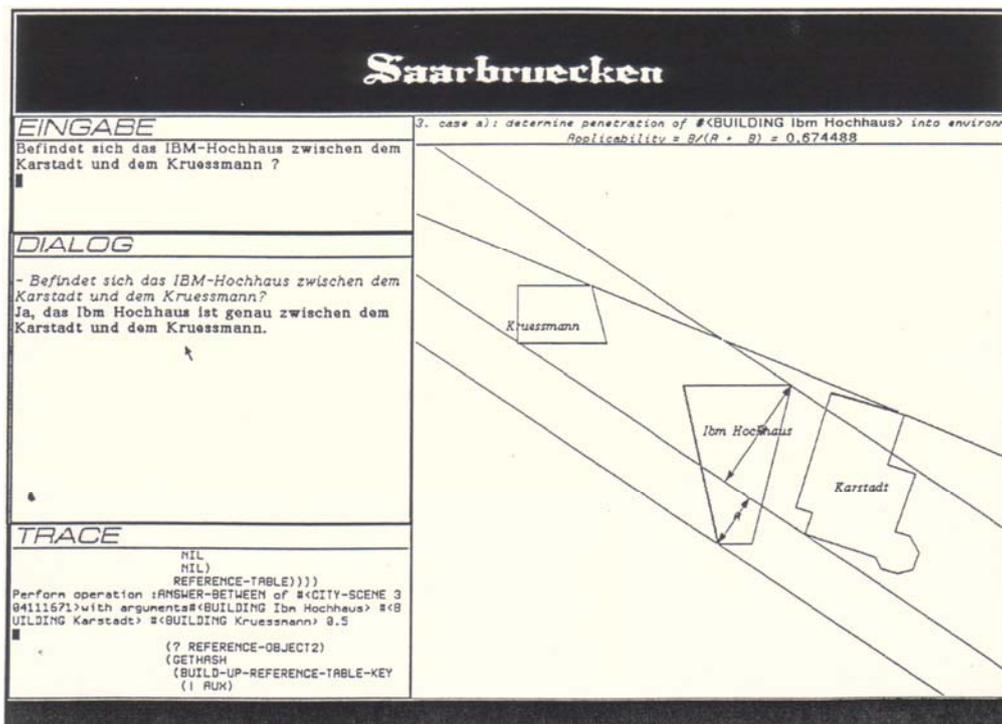


Abbildung 3

Die Unterscheidung zwischen Anwendbarkeit und Nichtanwendbarkeit eines sprachlichen Ausdrucks reicht nicht aus, um eine räumliche Situation adäquat zu beschreiben. Vielmehr müssen die Grenzen als fließend angesehen werden. Um dieser Tatsache Rechnung zu tragen, wird jedem Relationentupel ein Anwendbarkeitsgrad zugeordnet. In VITRA wird ein Wert aus dem reellen Intervall von 0 bis 1 verwendet, wobei 0 für nicht anwendbar und 1 für voll anwendbar steht. Veranschaulichen läßt sich die graduierte Anwendbarkeit durch Partitionierung eines Anwendbarkeitsraums in Regionen gleicher Anwendbarkeit, mit denen dann verschiedene Gradpartikel wie z.B. "ungefähr" oder "genau" assoziiert werden können.

Beispielsweise erfolgt die Berechnung des Anwendbarkeitsgrades der dreistelligen statistischen Relation "Zwischen" durch zwei Algorithmen: einerseits wird mit Hilfe der beiden Tangenten an den Bezugsobjekten die relative Eintauchtiefe des Relationssubjektes in den Zwischenraum berechnet (vgl. Abb. 3). Andererseits wird der Winkel $\langle \alpha \rangle$ zwischen den beiden Geraden berechnet, die die beiden Schwerpunkte der Referenzobjekte jeweils mit dem Schwerpunkt des Subjekts verbinden. Die Funktion $(\langle \text{Winkel} \rangle / 3.1416)$ ergibt dann einen Wert aus $[0 .. 1]$ (vgl. Abb. 4). Das Maximum der beiden berechneten Werte wird als Anwendbarkeitsgrad der Relation verwendet.

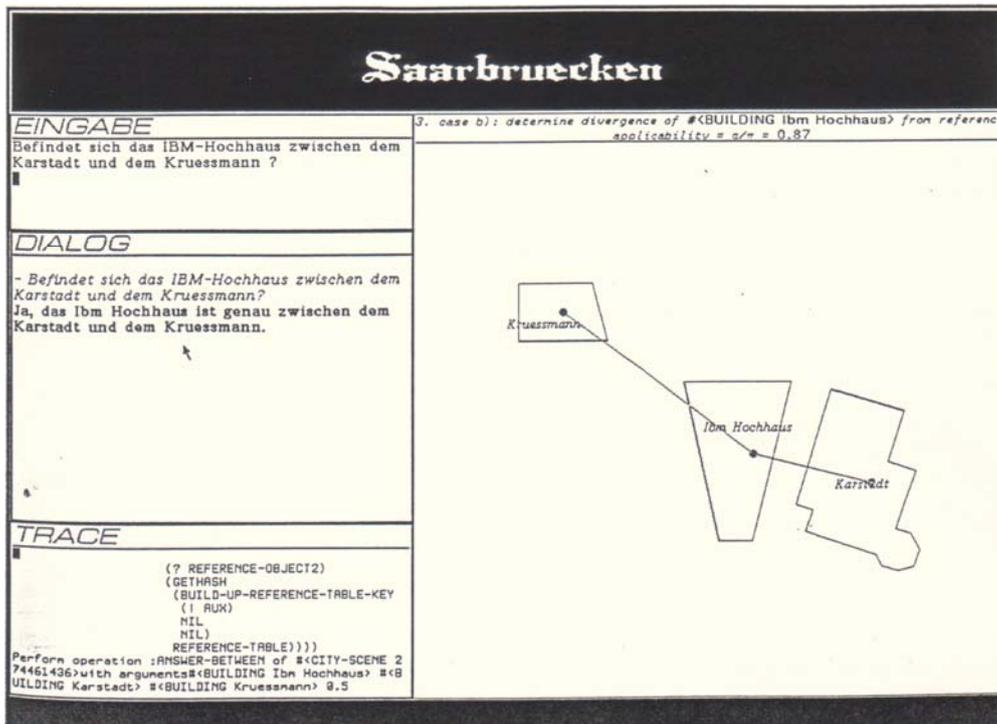


Abbildung 4

4. Anwendungen von Wegauskunftssystemen in Fahrzeugen

Das bereits seit zwei Jahren kommerziell angebotene System Travelpilot IDS von Bosch hat auf einem CD-Speicher die für einen Fahrer wichtigsten Katasteramtsdaten als Polygonzüge mit Straßennamen gespeichert: 280 000 km Überlandstraßen, 82 Groß- und Mittelstädte und 25 000 kleinere Orte. Nach einer einmaligen Normierung des Systems auf den momentanen Standort des Fahrzeugs verfolgt ein Bordrechner mithilfe von Sensoren an den Rädern und einem angeschlossenen Kompaß die Fahrtstrecke, ohne daß dafür spezielle Vorkehrungen auf den Fahrwegen (z.B. Induktionsschleifen, Sender) oder eine Satelliten-Navigation nötig wären. Auf einem neben dem Lenkrad angebrachten kleinen Bildschirm bekommt der Fahrer während der Fahrt jeweils seine Position, sein zu Beginn der Fahrt graphisch eingegebenes Ziel und den relevanten Ausschnitt des Straßennetzes graphisch angezeigt, wobei er zwischen neun verschiedenen Maßstäben wählen kann. Es ist offensichtlich, daß der Travelpilot durch die automatische Positionsanzeige und die Wahl des relevanten Ausschnitts der digitalen Karte eine wesentliche Verbesserung gegenüber dem manuellen Suchen in konventionellen Straßenkarten darstellt. Trotzdem ist die Leistung des Systems bei weitem nicht mit der eines guten Beifahrers zu vergleichen. Besonders störend wirkt der ständige visuelle Abgleich zwischen der Bildschirmdarstellung und der wahrgenommenen Umgebung. Wenn, wie derzeit geplant, die Datensätze des Travelpilot-Systems um Landmarken ergänzt werden, kann mit dem VITRA-System über den bisher verwendeten geometrischen Daten eine sprachliche Wegbeschreibung erstellt werden, die dann inkrementell lautsprachlich realisiert wird. Damit kann der Fahrer seinen Blick konstant auf die Fahrbahn richten und sich voll auf das Steuern des Fahrzeugs konzentrieren.

Nachteilig bei dem derzeitigen System ist noch, daß keine Information über den globalen Straßenzustand verfügbar ist, so daß eine Verbindung über einen Feldweg gleichgesetzt wird mit der über eine Bundesstraße. Durch das Vorhalten der gesamten kartographischen Information im Fahrzeug ist es auch nicht möglich, den aktuellen Straßenzustand (z.B. Unfall, Bauarbeiten, Stau) bei der Wegsuche zu berücksichtigen. Hier setzt ein in Entwicklung befindliches System von Motorola an, daß die gesamte Verkehrssituation berücksichtigt, und dem Fahrer automatisch erzeugte Wegbeschreibungen per Autotelephon zugänglich machen soll. Auch hier bieten die im VITRA-Projekt entwickelten Verfahren eine wichtige Grundlage für die Realisierung des anwendungsreifen Systems. Ein solches System stellt nicht nur eine Hilfe für den einzelnen Fahrer dar, sondern kann langfristig auch zu einer verbesserten Verkehrssituation führen, weil Einzelfahrzeuge individuell umgelenkt

werden können, wodurch bei Engpässen eine beliebige Entzerrung des Verkehrs auch über mehrere parallele, nicht ausgeschilderte Umleitungen erfolgen könnte.

Das Beispiel der automatischen Wegauskunft zeigt auch, daß für sinnvolle Anwendungen häufig deterministische und optimale Verfahren, wie sie bei der Suche nach effizienten Algorithmen entdeckt wurden, mit wissensbasierten Methoden, wie sie im Forschungsgebiet Künstliche Intelligenz entwickelt werden, kombiniert werden müssen, um eine für den Menschen wirklich nützliche maschinelle Problemlösung zu finden. Die Verwendung von Landmarken, die Einführung einer gewissen Redundanz der Beschreibung, die explizite Berücksichtigung des vermuteten Vorwissens des Fragenden und die Suche nach einer möglichst einfachen und gut merkbaren Auskunft sind Merkmale eines Ansatzes, der die Informationsverarbeitungsleistung des menschlichen Benutzers beim Systementwurf berücksichtigt. Dies entspricht dem Trend in der Informatik, daß der anteilige Entwicklungsaufwand für intelligente Benutzerschnittstellen im Vergleich zu den einer Anwendung zugrundeliegenden Berechnungsverfahren immer größer wird.

Literatur

- G. Herzog, C.-K. Sung, E. Andre, W. Enkelmann, H.-H. Nagel, T. Rist, W. Wahlster und G. Zimmermann (1989): Incremental Natural Language Description of Dynamic Imagery. In: Ch. Freksa und W. Brauer (Hrsg.), Wissensbasierte Systeme. 3. Internationaler GI-Kongreß, S. 153-162, Berlin, Heidelberg: Springer.
- W. Hoepfner, M. Carstensen und U. Rhein (1989): Wegauskünfte; Die Interdependenz von Such- und Beschreibungsprozessen. In: Ch. Freksa und Ch. Habel (Hrsg.), Repräsentation und Verarbeitung räumlichen Wissens, Berlin, Heidelberg: Springer.
- S. Müller (1989): CITYGUIDE - Wegauskünfte aus dem Computer. Memo 37, Universität des Saarlandes, SFB 314 .
- G. Retz-Schmidt (1988): Various Views on Spatial Prepositions. In: AI Magazine, Vol. 9, No. 2, 95-105.
- W. Wahlster (1989): One Word Says More Than a Thousand Pictures. On the Automatic Verbalization of the Results of Image Sequence Analysis Systems. In: Computers and Artificial Intelligence Journal, Vol. 8., No. 5, 479-492.