
Graphenbasierte Navigation eines Geometrischen Agenten: Integration von Perzeption und Instruktion

Diplomarbeit am Arbeitsbereich WSV
des Fachbereichs Informatik
der Universität Hamburg

von

Jan Hendrik Helwich
Kantstraße 36
22089 Hamburg

Betreuer: Prof. Dr. Christopher Habel
Zweitbetreuerin: Hedda Rahel Schmidtke

Dezember 2003

Inhaltsverzeichnis

1	Einleitung.....	5
1.1	Mentale und räumliche Repräsentationen.....	6
1.1.1	Kognitive Karten.....	6
1.1.2	Landmarken und Routenwissen.....	10
1.1.3	Routengraphen.....	11
1.1.4	Kognitive Modelle.....	12
1.2	Vergleich und Integration räumlicher Repräsentationen.....	14
1.2.1	Semantische Ähnlichkeit.....	14
1.2.2	Koreferenz.....	15
1.3	Aufbau der Arbeit.....	17
2	Der Geometrische Agent.....	18
2.1	Agentensysteme.....	18
2.2	Instruktionsphase.....	22
2.3	Die Umgebung des Agenten.....	24
2.4	Navigationsphase.....	27
3	Grundlagen der Wissensrepräsentation.....	30
3.1	Referentielle Netze.....	30
3.1.1	Die Repräsentationssprache SRL.....	31
3.1.2	Referentielle Netze.....	37
3.2	Taxonomien.....	39
3.3	CRIL-Graphen.....	45
4	Ähnlichkeitsmaße.....	48
4.1	Ähnlichkeit von Repräsentationen.....	49
4.2	Ähnlichkeit von Konzepten.....	51
4.3	Ähnlichkeit von Knoten.....	56
4.4	Ähnlichkeit von Graphen.....	61
4.5	Berechnung der Ähnlichkeit zweier Graphen.....	64
5	Repräsentation von Routenwissen.....	69
5.1	Routen.....	69
5.2	Routen in Wegbeschreibungen.....	71
5.2.1	Plätze.....	72
5.2.2	Pfade und Routensegmente.....	74
5.3	Räumliches Inventar der CRIL.....	76
5.3.1	Sorten.....	77
5.3.2	Relationen.....	78
5.4	Koreferenz von Wegen der Instruktion und der Perzeption.....	81

6	Koreferenz in der Navigationsphase.....	86
6.1	Berechnung der Koreferenzen während der Navigation.....	86
6.2	Berechnung der relevanten Teile des Instruktionsgraphen.....	91
6.3	Einfügen der Koreferenzkanten und Konfidenzwertberechnung	95
6.4	Ein Beispiel.....	96
7	Ausblick.....	101
8	Literaturverzeichnis	103

Kapitel 1

Einleitung

Für die Bewegung eines mobilen Agenten, wie einem Mensch, einem Tier oder einem Roboter, von einem Ort in seiner Umgebung zu einem anderen werden besondere Fähigkeiten benötigt. Zu diesen Fähigkeiten gehören zum Beispiel das Erkennen und Erinnern von Objekten, die Fähigkeit der Rückkehr zu bestimmten Orten oder das Lernen von Routen von einem Ort zu einem anderen. Die Verwendung dieser Fähigkeiten für die Bewegung lässt sich unter dem Begriff Navigation zusammenfassen. Einfache Navigationsaufgaben wie Futtersuche oder Pfadintegration¹ können ohne mentale Repräsentationen der Umgebung gelöst werden. Für komplexere Aufgaben wird eine mentale Repräsentation der Umgebung benötigt. Ein Beispiel für eine komplexere Navigationsaufgabe ist die Navigation eines Menschen mit Hilfe einer mündlich und im Voraus gegebenen Wegbeschreibung. Um zu navigieren, muss der Mensch sich die Wegbeschreibung gemerkt haben, also eine mentale Repräsentation der Wegbeschreibung erstellt haben, die er dann zur Navigation verwenden kann.

Ein Agent, der mit Hilfe einer Wegbeschreibung navigieren soll, hat verschiedene kognitive Aufgaben zu lösen. Zunächst muss er die Wegbeschreibung (Instruktion²) mental oder (im Fall eines künstlichen Agenten) intern repräsentieren. Dies setzt zunächst Sprachverstehen sowie eine adäquate Repräsentation des erhaltenen räumlichen Wissens voraus. Während der Navigation muss der Agent wahrnehmen und das aus der Wahrnehmung (Perzeption) stammende räumliche Wissen mit der mentalen bzw. internen Repräsentation vergleichen, um die in der Instruktion beschriebenen Objekte und räumliche Konfigurationen in der Umgebung zu identifizieren. Da Wegbeschreibungen weder alle Objekte und räumlichen Konfigurationen entlang des beschriebenen Weges enthalten (vgl. [TvL99]) noch jedes Objekt vollständig beschrieben ist, muss der autonom navigierende Agent (im Folgenden Navigator genannt) die Fähigkeit haben, die Ähnlichkeit zwischen beschriebenen und wahrgenommenen Objekten zu bestimmen.

¹ Pfadintegration ermöglicht es einem Agenten, nach einer Anzahl von Bewegungen und Richtungswechseln den Startort der Bewegungen, zum Beispiel das Nest, zu bestimmen.

² Im Folgenden werden neu eingeführte Begriffe stets durch arialen Font dargestellt.

Der in [TSH03] vorgestellte geometrische Agent soll mit Hilfe einer im Voraus gegebenen Wegbeschreibung navigieren. Die in der Wegbeschreibung enthaltene Information ist in der Conceptual Route Instruction Language (CRIL) repräsentiert. Die räumlichen Informationen aus Instruktion und Perzeption sind als graphenartige Strukturen (CRIL-Graphen) repräsentiert. Knoten und Kanten der Graphen repräsentieren linguistische und räumliche Konzepte. In der vorliegenden Arbeit soll gezeigt werden, wie der aus der Instruktion stammenden CRIL-Graph mit den aus der Perzeption stammenden CRIL-Graphen verglichen und ein Maß der Ähnlichkeit für CRIL-Graphen bestimmt werden kann. Dabei wird konzeptuelles Wissen über die repräsentierten Konzepte verwendet. Weiterhin wird ein Vorschlag gemacht, wie der für den Vergleich relevante Teilgraph im aus der Instruktion stammenden CRIL-Graphen gefunden werden kann, wie damit eine Selektion der Perzeption erreicht werden kann und das Maß der Ähnlichkeit verwendet werden kann, um die in der Instruktion beschriebenen Objekte in der Umgebung des Agenten zu identifizieren.

Die folgenden Abschnitte sollen zunächst die vorliegende Arbeit in den Bereich der Kognitionswissenschaft einordnen sowie einige Grundlagen und Begriffe der räumlichen Repräsentation und der Bestimmung von Ähnlichkeiten zwischen Repräsentationen erörtern.

1.1 Mentale und räumliche Repräsentationen

Die vorliegende Arbeit soll den Vergleich und die Integration von CRIL-Graphen im von Tschander et al. in [TSH03] vorgestellten geometrischen Agenten ermöglichen. CRIL-Graphen sind kognitive Modelle mentaler räumlicher Repräsentationen. Ein in der Kognitionswissenschaft bekanntes Modell mentaler räumlicher Repräsentationen sind kognitive Karten. Auch CRIL-Graphen können als kognitive Karten verstanden werden. Um dies zu zeigen, wird im Folgenden das Modell der kognitiven Karten kurz eingeführt (Abschnitt 1.1) und anschließend deutlich gemacht, dass auch bestimmte Routengraphen als Modell für kognitive Karten verwendet werden können (Abschnitt 1.2). Abschnitt 1.3 soll einen Überblick über die kognitive Modellierung und die Rolle der Informatik in der Kognitionswissenschaft geben.

1.1.1 Kognitive Karten

In der Forschung der Repräsentation räumlichen Wissens von Menschen und Tieren hat das Modell der kognitiven Karten in den letzten Jahren stark an Popularität gewonnen. Das Konzept der kognitiven Karten wurde in der Psychologie entwickelt, um Verhaltensweisen von Tieren und Menschen zu erklären, die nur durch das Vorhandensein einer internen Repräsentation der

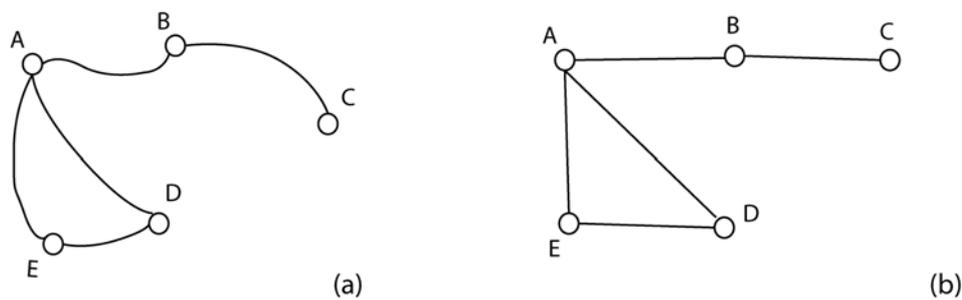


Abbildung 1.1: (a) Ein Netzwerk von Plätzen einer Umgebung. Die Plätze sind durch vom Navigator benutzbare Wege verbunden. Die Wegführung kann z. B. durch Hindernisse entstehen. (b) Die dazugehörige Abbildung auf einen Graphen, in dem die Erreichbarkeit der Plätze untereinander durch die Kanten dargestellt ist.

Umgebung erklärt werden können. Ein Beispiel ist die Fähigkeit einiger Tiere und des Menschen, Abkürzungen in Umgebungen zu finden, in denen keine markanten Objekte sowohl vom Startort der Navigation als auch vom Zielort aus sichtbar sind. Deshalb ist Annahme gerechtfertigt, dass (mindestens) die in einer kognitiven Karte enthaltene Information auch in einer mentalen räumlichen Repräsentation des Menschen enthalten sind.

Zunächst soll deutlich werden, welche Funktionen kognitive Karten in der Raumkognition haben. Poucet unterscheidet in [Pou93] S.164 drei Funktionen von kognitiven Karten in Tieren:

„Broadly speaking, three distinct functions can be defined as (a) spatial recognition and identification, (b) spatial localization and memory, and (c) planning of spatial actions.“

Dabei weist Poucet darauf hin, dass die Funktionen nicht unabhängig voneinander sind, da sie integrale Bestandteile des gleichen Systems sind. So kann ein Tier oder ein Mensch seine Position nicht bestimmen, ohne seine Umgebung zu identifizieren und es bzw. er kann seine Aktionen nicht planen, ohne seine Position bestimmt zu haben.

Die grundlegenden Elemente der kognitiven Karten sind Repräsentationen von Plätzen und Relationen, die den topologischen Zusammenhang der Plätze beschreiben. Plätze sind Orte der Umgebung. Die Relationen repräsentieren Wege, welche dem Navigator bekannt sind und welche die in Relation stehenden Plätze verbinden. Ohne zunächst auf die Repräsentation von Plätzen einzugehen, kann an dieser Stelle bereits festgehalten werden, dass kognitive Karten somit eine graphenartige Struktur haben. Dabei sind die Platzrepräsentationen die Knoten und die zwischen ihnen geltenden topologischen Relationen die Kanten des Graphen. Abbildung 1.1 soll dies anhand einer schematischen Umgebung und einem dazugehörigen Graphen, der die Erreichbarkeit der Plätze untereinander repräsentiert, verdeutlichen.

Bei der Repräsentation von Plätzen lassen sich zwei Ansätze unterscheiden. Der häufig für kognitive Karten von Tieren verwendete Ansatz ist, dass ein Platz durch eine Sicht repräsentiert ist. Die Sicht entspricht dabei dem vom

Navigator wahrgenommenen Bild beim Erreichen des Platzes. Entspricht während der Navigation das wahrgenommene Bild der repräsentierten Sicht, so ist der Platz identifiziert. Die darauf folgende Bewegung kann mit der Sicht verknüpft sein, was zu der so genannten „place recognition-triggered response“ führt (vgl. z. B. [TWB97] S. 499). Der Ansatz der sichtbasierten Repräsentation von Plätzen hat aber verschiedene Schwächen. Zum einen lässt sich mit Sicherheit sagen, dass der Mensch nicht nur Sichten für die Repräsentation von Plätzen verwendet. Ansonsten ließe sich die unbestritten gute Navigationsleistung von Menschen mit Hilfe einer mündlichen Wegbeschreibung (in der keine Bilder übermittelt werden) nicht erklären. Verschiedene Experimente mit Tieren haben außerdem gezeigt, dass auch Tiere andersartige Repräsentationen von Plätzen zu haben scheinen. Dieser Schluss scheint gerechtfertigt, da es einigen Tieren möglich ist, einen Platz unabhängig von einer speziellen Sicht zu identifizieren (vgl. z. B. [Pou93] S.169ff). Um zu gewährleisten, dass eine Platzrepräsentation unabhängig von einer speziellen Sicht ist, muss der Platz relativ zu seiner Umgebung oder durch seine intrinsischen Charakteristika (so genannte local cues) repräsentiert werden. Die Verwendung von local cues alleine ist aber immer noch nicht ausreichend, um die Erkennungsleistung von Plätzen zu erklären, die einige Tiere erbringen. Morris stellt in [Mor81] ein Experiment vor, in dem ein Tier eine sichere, aber unter Wasser verborgene Plattform in einem Wasserlabyrinth erreichen soll. Die Tatsache, dass die Versuchstiere es nach wenigen Versuchen schaffen, auf relativ direktem Weg die Plattform zu erreichen, lässt den Schluss zu, dass nicht nur local cues für die mentale Repräsentation von Plätzen verwendet werden. Daraus lässt sich folgender aus [Pou93] S.165 stammender Schluss ziehen:

„As a consequence, a place representation should encode, among other things, sufficient information about the spatial relationships between that place and the environment.“

Zusätzlich zu den Informationen, die zur Identifizierung der Plätze benötigt werden, müssen Informationen über die Wege zu anderen Plätzen repräsentiert sein. Die Information, welche Plätze von einem Platz aus erreichbar sind, ist bereits durch die topologischen Relationen gegeben. Dies reicht aber nicht aus, um an einem Platz zu entscheiden, welchen der von dem Platz abgehenden Wege der Navigator nehmen muss, um zu einem von dort aus direkt erreichbaren Platz zu kommen. Um dies entscheiden zu können, müssen die räumliche Lokalisation der Wege repräsentiert werden, also zum Beispiel, dass ein Weg links von einem markanten Objekt der Umgebung vorbei läuft. Ein möglicher Ansatz solcher Repräsentationen ist, die räumliche Objektlokalisierung durch Vektoren, also Polarkoordinaten zu kodieren (vgl. z. B. [OKe91]). In Wegbeschreibungen hingegen werden meist Anordnungsrelationen oder topologische Relationen verwendet, die zum Beispiel durch Präpositionen wie *vor*, *über*, *in* oder *rechts* repräsentiert sind. Dies deutet darauf hin, dass räumliche Relationen für den Menschen hinreichend durch die

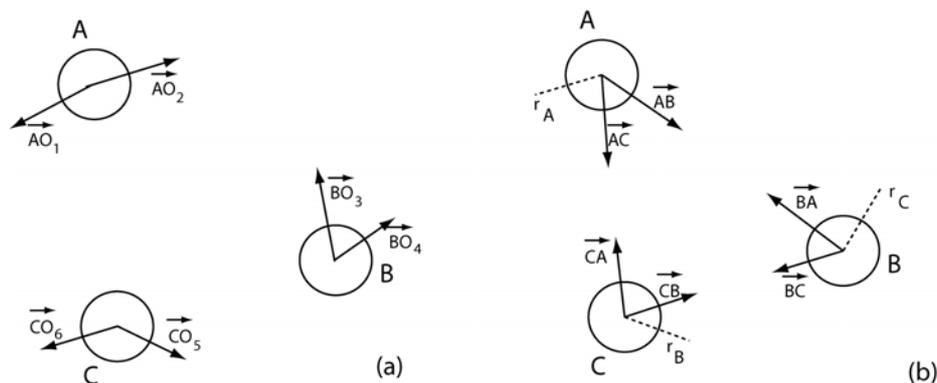


Abbildung 1.2: Ein Beispiel für die schematische Repräsentation von Plätzen A, B und C. (a) Vektoren von den Plätzen zu Objekten O₁ bis O₆ der Umgebung für die Identifikation des Platzes. (b) Vektoren zur Bestimmung der räumlichen Lage von Wegen zu anderen Plätzen. Jeder Platz hat eine eigene Referenzrichtung r_A , r_B und r_C .

sprachlichen (qualitativen) Konzepte der eben beschriebenen Relationen repräsentiert sind.

Sowohl die Verwendung von Vektoren als auch die Verwendung von (einigen) sprachlichen räumlichen Relationen erfordert eine Referenzrichtung (vgl. zum Beispiel [HeS98]). Eine globale Referenzrichtung kann nicht angenommen werden, es sei denn, das Tier oder der Mensch verfügt über eine Art internen Kompass. Es kann also angenommen werden, dass Tiere oder Menschen lokale Referenzrichtungen verwenden (vgl. z.B. [Pou93], [TWB97]). Die lokale Referenzrichtung kann unterschiedlich repräsentiert sein. Als Beispiel soll an dieser Stelle die Verwendung einer Sicht genügen. Der Navigator kann sich dann an dem Platz in eine Position bringen, in der das momentan wahrgenommene Bild mit der repräsentierten Sicht übereinstimmt, und hat sich selbst somit entsprechend der lokalen Referenzrichtung ausgerichtet. Abbildung 1.2 zeigt eine schematische Repräsentation von Plätzen, in der Vektoren sowohl für die Identifizierung der Plätze (a) als auch für die Repräsentation der räumlichen Lage der Wege (b) verwendet werden.

Es lässt sich zusammenfassend festhalten, dass kognitive Karten graphenartige Struktur haben. Knoten der Graphen repräsentieren Plätze. Kanten repräsentieren topologische Relationen, genauer die Erreichbarkeit der Plätze untereinander. Kanten repräsentieren somit Wege in der Umgebung, die der Navigator kennt. An die Repräsentation der Plätze wird folgende Forderung gestellt:

Die Repräsentationen der Plätze muss es dem Navigator ermöglichen,

- (a) den Platz zu identifizieren, was zum Beispiel mit Hilfe von Informationen über räumliche Relationen des Platzes zu anderen Objekten der Umgebung oder local cues ermöglicht wird und
- (b) zu entscheiden, wie er einen bestimmten Platz von dem identifizierten Platz aus erreichen kann. Dies wird zum Beispiel erreicht, indem die räumliche Lage der abgehenden Wege repräsentiert wird.

1.1.2 Landmarken und Routenwissen

Während der Navigation scheinen markante Objekte der Umgebung von entscheidender Bedeutung zu sein. Sie werden zum Beispiel zur Identifikation von Plätzen oder der Lokalisierung von Objekten vom Navigator verwendet. Solche markanten Objekte der Umgebung werden in der Literatur meist **Landmarken** genannt. Es sollte bereits deutlich geworden sein, dass Landmarken im räumlichen Wissen von entscheidender Bedeutung sind. Untersuchungen von Wegbeschreibungen wie zum Beispiel in [Hab88], [Kle79], [HeS98] und [DaD98] bestätigen diese These. Der Begriff Landmarke wird von vielen Autoren auf [Lyn73] zurückgeführt. Deshalb soll der Begriff Landmarke durch folgendes Zitat aus [Lyn73] S. 48 bestimmt sein:

„Landmarks are another type of point-reference, ... They are usually a rather simply defined physical object: building, sign, store, or mountain. ... Some landmarks are distant ones, typically seen from many angles and distances, over the tops of smaller elements, and used as radial references. ... Other landmarks are primarily local, being visible only in restricted localities and from certain approaches.”

Teilweise wird zwischen Weg- und Landmarken unterschieden. Dabei sind Wegmarken Objekte, die in relativer Nähe zur Route liegen, Landmarken sind Objekte die in einiger Entfernung liegen. Diese Unterscheidung soll in der vorliegenden Arbeit nicht gemacht werden.

Eine **Route** soll im Folgenden eine Verbindung von zwei Orten A und B der Umgebung sein, die es einem Agenten ermöglicht von A nach B zu gelangen. In der Kognitionspsychologie wird häufig zwischen Landmarken-, Routen- und Überblickswissen unterschieden. Dabei stellt das Landmarkenwissen die erste Stufe des Erwerbs räumlichen Wissens durch Exploration der Umgebung dar. Landmarkenwissen besteht nur aus Repräsentationen von Landmarken in der Reihenfolge, in der sie entlang der explorierten Route liegen (vgl. z. B. [GiM97]). Routenwissen enthält zusätzlich Wissen darüber, welche Art von Reorientierung an den Landmarken vorgenommen werden soll und wie bestimmte Objekte der Umgebung lokalisiert sind (vgl. z. B. [WKH00]). Überblickswissen enthält die Repräsentationen mehrere Routen, die zu einer Repräsentation integriert wurden. Der Übergang von Routen- zu Überblickswissen erfordert vom Wissensträger die Fähigkeit zu schließen, dass es sich bei zwei Plätzen aus zwei verschiedenen Routen um ein und den selben Platz handelt. Die Repräsentationen der Routen werden dann an diesem Platz zusammengefügt. Die Übergänge während des Erwerbs von Wissen über die Umgebung sind (selbstverständlich) fließend und eine klare Abgrenzung zwischen den einzelnen Stadien ist nicht möglich. Es sei an dieser Stelle festgehalten, dass das durch Routeninstruktionen kommunizierte Wissen Routenwissen entspricht.

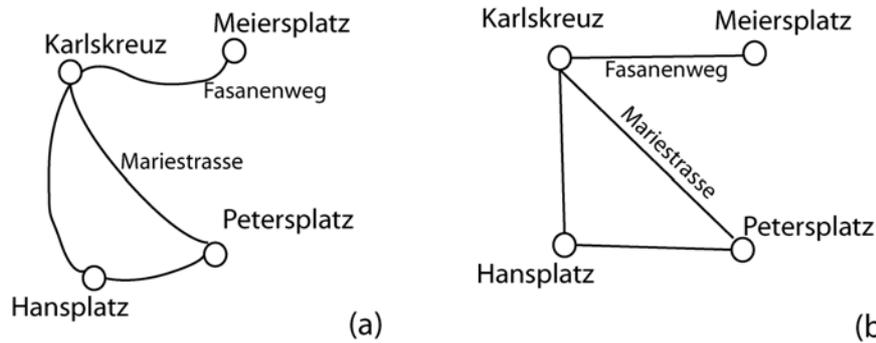


Abbildung 1.3: Eine Beispielumgebung (a) und der dazugehörige Routengraph (b), in dem die Namen der Plätze und Straßen Markierungen der Knoten und Kanten sind.

1.1.3 Routengraphen

Betrachtet man Abbildung 1.1 (b), so fällt auf, dass es sich bei dem dargestellten Graphen um einen Routengraphen handelt, wie er zum Beispiel für Streckennetzpläne der öffentlichen Verkehrsmittel verwendet wird. Streckennetzpläne erfüllen die in Abschnitt 1.1.1 geforderten Eigenschaften von kognitiven Karten für Menschen, die mit öffentlichen Verkehrsmitteln reisen. Knoten des Streckennetzes repräsentieren Haltestellen und somit Plätze der Umgebung. Die Erreichbarkeit der Plätze ist im Plan durch die Kanten dargestellt. Die Platzrepräsentationen enthalten die Namen der Plätze, und somit genug Information, um dem Mensch eine eindeutige Identifizierung des Platzes zu ermöglichen. Die Identifikation von Wegen zu einem erreichbaren Platz ist durch die Angabe einer Liniennummer o. ä. repräsentiert.

Für die Repräsentation des räumlichen Wissens, das in einer Wegbeschreibung enthalten ist, reicht dies noch nicht aus. Angenommen der Navigator erhält die folgende Wegebeschreibung:

- (1) „Wenn du am Karlskreuz stehst, gehst du zwischen der Brauerei und der Post durch. Dann kommst du direkt zum Hansplatz.“

Die Wegbeschreibung soll sich auf die in Abbildung 1.3 (a) dargestellte Umgebung beziehen. In Abbildung 1.3 (b) ist der dazugehörige Routengraph dargestellt. Der Routengraph ist offensichtlich nicht ausdrucksstark genug, um die in der Wegbeschreibung enthaltene Information über die räumliche Lage des Weges vom Karlskreuz zum Hansplatz zu repräsentieren. Es muss also eine Erweiterung des Graphen vorgenommen werden, die es ermöglicht, die räumlichen Informationen zu repräsentieren. Eine Möglichkeit ist, die Repräsentation der Knoten und Kanten durch Attribute und Prädikate zu erweitern. Dann ließe sich die in Beispiel (1) enthaltene räumliche Information durch den in Abbildung 1.4 (a) dargestellten Graphen repräsentieren. Die Erweiterung von Kanten durch Prädikate ist formell schwierig, da Kanten Relationen repräsentieren. Ersetzt man aber die Kanten durch Knoten, die in entsprechender Relation zu den Platz-Knoten stehen, ergibt sich eine äquivalente Repräsentation, in der die Repräsentation der Wege beliebig mit

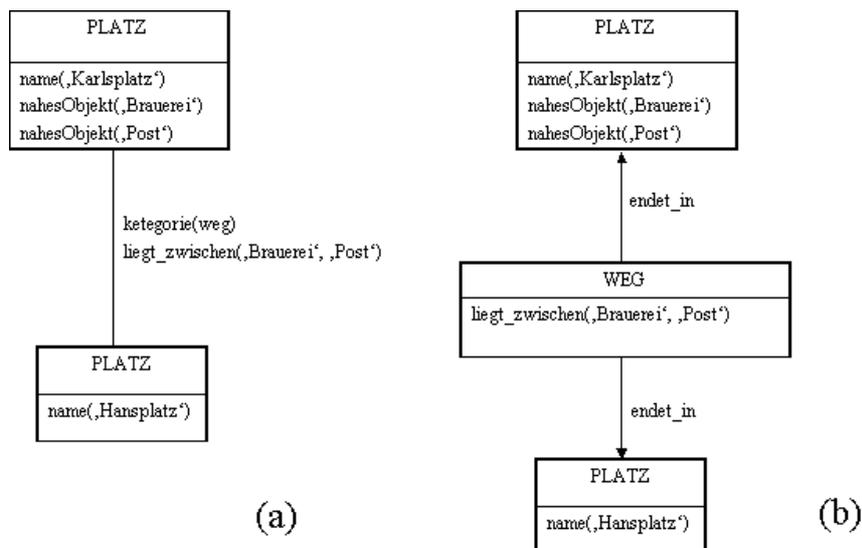


Abbildung 1.4: (a) Ein erweiterter Routengraph mit Prädikaten als Knoten- und Kantenbeschriftungen. Die Kategorien der Knoten sind in Großbuchstaben angegeben. (b) Ein Graph, in dem die Kante aus (a) durch einen Knoten ersetzt wurde. Die Kantenbeschriftung symbolisiert die durch die Kante repräsentierte Relation.

Prädikaten beschrieben werden kann. So ergibt sich der Graph in Abbildung 1.4 (b), der, wie sich noch zeigen wird, in der Struktur den CRIL-Graphen gleicht. Die Beschriftungen der Kanten in Abbildung 1.4 (b) beschreiben die Relationen, die sie repräsentieren. Die in Abbildung 1.4 dargestellten Prädikate sind nach ihrer Bedeutung benannt.³ Der in Abbildung 1.4 (b) dargestellte Graph entspricht in der Struktur den in Abschnitt 3.3 vorgestellten CRIL-Graphen. Mit den in Kapitel 5 vorgestellten Konzepten zur Repräsentation von Routenwissen können solche Graphen als ein Modell kognitiver Karten verwendet werden.

1.1.4 Kognitive Modelle

In den Kognitionswissenschaften, zum Beispiel der Kognitionspsychologie und der Psycholinguistik, wird nach Erklärungen für Verhalten und Fähigkeiten von Mensch und Tier gesucht. Die Entwicklung von Modellen auf der Grundlage von empirischen Daten soll (wie in jeder Wissenschaft) zum Verstehen der dem beobachteten Verhalten zugrunde liegenden Prozesse und Strukturen beitragen. Die Ansicht, dass sich Kognition durch mathematische Modelle darstellen und somit auch simulieren lässt, ist weit verbreitet, wie auch Simon und Wallach in [SiW99] S.1 deutlich machen:

³ Es sei hier darauf hingewiesen, dass in diesem Abschnitt kein Wert auf formelle Korrektheit gelegt wurde, sondern lediglich gezeigt werden soll, welche Struktur ein Graph haben muss um den Anforderungen einer kognitiven Karte zu entsprechen. Eine formelle Definition von CRIL-Graphen erfolgt in Kapitel 3.

„Over the past four decades the view that human cognition can be conceptualized as computational processes operating on symbolic (and subsymbolic) representations has grown into the modelling approach which is one of the methodological hallmarks of cognitive science, ...“

Durch rechnergestützte Simulation von kognitiven Modellen können Schlüsse gezogen werden, ob das Modell die Strukturen und Prozesse korrekt abbildet. Die Aussagekräftigkeit solcher Simulationen ist aber begrenzt. Dass die simulierten Ergebnisse mit den Ergebnissen der Versuche (größtenteils) übereinstimmen, bedeutet nicht unbedingt, dass das erstellte Modell korrekt ist. Es bedeutet nur, dass das Modell nicht grundlegend falsch ist. Nur keine oder eine sehr geringe Übereinstimmung lässt den Schluss zu, dass das Modell nicht korrekt ist. Ob die simulierten Ergebnisse mit den Ergebnissen der Versuche übereinstimmen oder nicht, muss (zusätzlich) oft subjektiv beurteilt werden. Um die Vermutung der Übereinstimmung zu erhärten, können die Ergebnisse der Simulation für weitere empirische Studien verwendet werden. Zum Beispiel kann die Umgebung in der Simulation derart verändert werden, dass ein vergleichbarer empirischer Versuch durchgeführt werden kann. Stimmen die Ergebnisse des Versuchs mit den Ergebnissen der veränderten Simulation überein, spricht dies wieder dafür, dass das Modell nicht falsch ist. So kann die Vermutung, dass ein Modell korrekt ist, nach und nach erhärtet werden. Für eine eingehendere Betrachtung kognitiver Modellierung möchte ich auf [SiW99] verweisen. An dieser Stelle seien nur zwei wichtige Punkte bezüglich der kognitiven Modellierung festgehalten:

1. Ein kognitives Modell soll empirisch beobachtetes Verhalten von Tieren oder Menschen modellieren. Die Vermutung, dass die modellierten kognitiven Strukturen und Prozesse korrekt abgebildet sind, lässt sich aber nur durch die Zusammenarbeit der kognitionswissenschaftlichen Disziplinen und Methoden erhärten.
2. Die Simulation des Modells soll es ermöglichen
 - a. festzustellen, ob das Modell grundlegend falsch ist oder, wenn dies nicht der Fall ist,
 - b. die Vermutung der Korrektheit des Modells durch weitere, auf den Ergebnissen der Simulation beruhende Forschungen zu erhärten.

Es sollte deutlich geworden sein, dass die Simulation von kognitiven Modellen nur die Annahme der Korrektheit eines Modells erhärten kann. Außerdem sollte deutlich geworden sein, dass die (Qualität der) Forschung im Bereich der Kognitionswissenschaft in hohem Maße von der Zusammenarbeit der einzelnen Disziplinen abhängt. Die Aussagen dieses Abschnitts treffen natürlich auch auf den Geometrischen Agenten zu, da dieser ein kognitives Modell eines nach einer natürlichen Wegbeschreibung navigierenden Menschen ist.

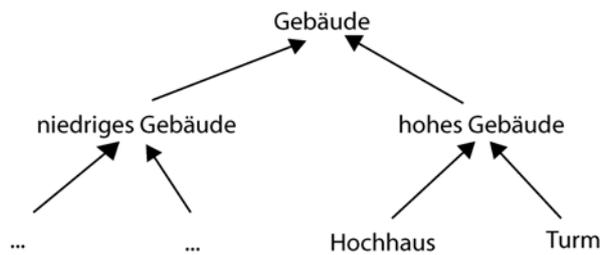


Abbildung 1.5: Ausschnitt einer Beispieltaxonomie. Die Pfeile entsprechen der Is-A-Beziehung.

1.2 Vergleich und Integration räumlicher Repräsentationen

Angenommen, ein Navigator erhält eine Wegbeschreibung, in der die Äußerung „an dem Hochhaus biegest du links ab“ enthalten ist. Während der Navigation nach der Wegbeschreibung kommt der Navigator an ein Gebäude, dass von ihm als ein Turm erkannt wird. Ein Mensch würde zumindest vermuten, dass es sich bei dem Turm um das beschriebene hohe Gebäude handelt, da ein Hochhaus nach menschlichem Empfinden ähnlich einem Turm ist. Aber worauf basiert die Ähnlichkeit von zwei Begriffen und wie kann in einer graphenartigen Repräsentation zum Ausdruck gebracht werden, dass zwei Knoten das gleiche Objekt repräsentieren? Diese beide Fragen sollen im Folgenden beleuchtet werden.

1.2.1 Semantische Ähnlichkeit

Ein verbreiteter Ansatz, die semantische, also inhaltliche Ähnlichkeit zweier Repräsentationen zu bestimmen, ist, den Anteil an Informationen⁴ zu vergleichen, der in beiden Repräsentationen enthalten ist (vgl. z. B. [ZZL02], [Res99]). Für die Bestimmung eines Ähnlichkeitsmaßes von zwei Repräsentationen soll in der vorliegenden Arbeit deshalb folgende informelle Forderung gelten:

- (1) Ein Ähnlichkeitsmaß zweier Repräsentationen soll sich auf die Informationen beziehen, die in beiden Repräsentationen enthalten sind.

Um ein Ähnlichkeitsmaß für zwei gegebene Repräsentationen zu bestimmen, muss also bekannt sein, welche Informationen in beiden Repräsentationen enthalten sind. Angenommen, die beiden Repräsentationen sind Knoten eines

⁴ Zu erklären was genau Information ist, bedarf einer eingehenden informationswissenschaftlichen Abhandlung und kann und soll in dieser Arbeit nicht geklärt werden. Im Folgenden werde ich davon ausgehen, dass eine Repräsentation Informationen enthält, was zum Ausdruck bringen soll, dass es sich bei der Information einer Repräsentation um eine Menge handelt.

Graphen wie er in Abschnitt 1.1.2 beschrieben wird. Die Knoten repräsentieren Objekte der Umgebung und die Eigenschaften der Objekte sind durch Attribute und Prädikate repräsentiert. Durch Vergleichen der Attribute und Prädikate der Knoten lässt sich leicht feststellen, welche Attribute und Prädikate in beiden Knoten gleich sind. Damit ist auch bekannt, welche Informationen in beiden Knoten enthalten sind. Nun kann es aber durchaus sein, dass Attribute oder Prädikate ihrerseits Begriffe repräsentieren, die ähnlich zueinander sind. Es muss also möglich sein, die Information zu bestimmen, die zwei durch Attribute oder Designationen repräsentierte Konzepte gemeinsam haben.

Eine Möglichkeit, diese Information zu bestimmen, ist, die Konzepte in so genannte Is-A-Taxonomien (auch Is-A-Hierarchien genannt) einzuordnen. In diesen Taxonomien wird die Zugehörigkeit von Konzepten zu allgemeineren Konzepten dargestellt. So kann die Aussage, dass Turm und Hochhaus hohe Gebäude sind, durch die Taxonomie in Abbildung 1.5 dargestellt werden. Dabei lässt sich sofort feststellen, dass Turm und Hochhaus die Information, ein hohes Gebäude zu sein, gemein haben. Wie Taxonomien verwendet werden können, um ein Maß der Ähnlichkeit zwischen einzelnen Konzepten zu bestimmen und damit ein Maß der Ähnlichkeit zwischen ganzen Graphen bestimmt werden kann, wird in Kapitel 4 erläutert.

Es soll nicht unerwähnt bleiben, dass die Ähnlichkeit zweier Konzepte kontextabhängig ist. Beim Vergleich zwischen dem Konzept Hochhaus und dem Konzept Turm lässt sich intuitiv feststellen, dass beide Konzepte für Gebäude stehen, die normalerweise sehr hoch sind. Die beiden Eigenschaften, ein Gebäude zu sein und hoch zu sein, sind im Kontext einer räumlichen Beschreibung wichtige Eigenschaften, da sie gut zu erkennende Objekte der Welt charakterisieren. Das spricht für eine hohe Ähnlichkeit der Konzepte im Kontext einer räumlichen Beschreibung. Ein anderer Kontext kann zu einer anderen Einschätzung der Ähnlichkeit führen. Ein Beispiel ist die Beschreibung von Architekturen. In diesem Kontext können die Eigenschaften, dass ein Turm normalerweise rund ist und wenige Fenster hat, wichtig erscheinen. Ein Hochhaus ist in den meisten Fällen aber eckig und hat viele Fenster. Aufgrund dieser unterschiedlichen Eigenschaften kann die Ähnlichkeit zwischen Turm und Hochhaus im Kontext der Architektur als gering angesehen werden. Da der Kontext während der Navigation eines autonomen Agenten feststeht, soll die Kontextabhängigkeit der Ähnlichkeit in der vorliegenden Arbeit nicht weiter behandelt werden. Einen Ansatz der Bestimmung eines Ähnlichkeitsmaßes, das den Kontext in Betracht zieht, wird zum Beispiel in [KMI95] vorgestellt.

1.2.2 Koreferenz

Es stellt sich die Frage, wie in einem Wissen repräsentierenden Graphen dargestellt werden kann, dass zwei Knoten das gleiche Objekt (der realen Welt oder des Denkens) repräsentieren. Intuitiv ließe sich annehmen, dass die beiden Knoten gleich sind. Diese Annahme würde dazu führen, dass es in der Menge

der Knoten nur einen, anstatt der zwei Knoten gibt. In der Implementation würden die beiden Knoten zu einem Knoten verschmolzen werden. Das ist nicht unbedingt korrekt und wünschenswert. Statt dessen bietet es sich an, diese Art der Gleichheit explizit durch eine Kante im Graphen zu repräsentieren. Die aus der Wissensrepräsentation bekannte Koreferenz drückt eben diese Gleichheit der repräsentierten Objekte aus.

Der Begriff der Extension eines Konzepts ist aus der Wissensrepräsentation und Psychologie bekannt und beschreibt die Menge der Objekte, die durch das Konzept repräsentiert werden. Im Folgenden soll deshalb die Extension eines CRIL-Knotens die Menge der Objekte sein, die durch ihn repräsentiert sind. Verschiedene Repräsentationen werden koreferent genannt, wenn sie die gleiche Extension haben. Reimers definiert zum Beispiel die Koreferenzbeziehung von Knoten in semantischen Netzen in [Rei91] auf S.119 folgendermaßen.

„Eine Koreferenzkante zwischen zwei Konzeptknoten (die für Individualkonzepte oder Konzeptklassen stehen können) postuliert, dass die (nicht leeren) Extensionen der durch sie repräsentierten Konzepte gleich sind. ...“

Diese Definition gelte auch für die Koreferenzkante in CRIL-Graphen, wobei die repräsentierten Konzepte die Objekte des Denkens oder der Welt sind. Wegen der Annahme einer geschlossenen Welt (closed world assumption) sollen Knoten, die nicht durch eine Koreferenzkante verbunden sind, als nicht koreferent interpretiert werden. Für die Bestimmung der Koreferenz im Geometrischen Agenten muss allerdings beachtet werden, dass es sich in vielen Fällen nur um ein vom Agenten vermutete Koreferenz handelt, da die Koreferenz mithilfe der Ähnlichkeit ermittelt wurde. Dies muss auch explizit in der Repräsentation der Koreferenzkante vermerkt werden. Deshalb wird für jede Koreferenzkante ein Konfidenzwert ermittelt, der angibt, wie sicher es ist, dass die Kante eine korrekte Koreferenz darstellt.

Die Verschmelzung von aus der Wegbeschreibung stammenden Knoten mit aus der Perzeption stammenden Knoten ist deshalb nicht wünschenswert, da während der Navigation erkannte Koreferenzen nicht unbedingt korrekt sind. Zum Beispiel ist ein Navigator nicht unbedingt in der Lage, alle Objekte der Umgebung wahrzunehmen. So kann es passieren, dass zunächst ein Objekt wahrgenommen wird, das anscheinend einem Objekt der Wegbeschreibung entspricht. Zu einem späteren Zeitpunkt wird aber ein Objekt in der Umgebung wahrgenommen, das dem beschriebenen Objekt besser entspricht. In diesem Fall muss der Navigator in der Lage sein, seine Annahme zu revidieren. Es sollte daher möglich sein, als koreferent repräsentierte Knoten wieder als verschiedene Knoten zu repräsentieren. Nur dann kann nämlich die neue Koreferenz repräsentiert werden.

Falsch ist die Annahme, dass zwei Repräsentationen mit der gleichen Extension gleich sind, insbesondere dann, wenn die Repräsentationen zu verschiedenen Kontexten gehören und die Repräsentationen des einen

Kontextes von außen, also aus einem anderen Kontext nicht referenziert werden darf. So darf zum Beispiel die Repräsentation der Wegbeschreibung aus dem Kontext der Perzeption nicht direkt referenziert werden. Informell würde eine solche direkte Referenz bedeuten, dass der Navigator sich nicht mehr der Tatsache bewusst ist, dass die Repräsentationen aus unterschiedlichen Quellen stammen. Man spricht von opaken Kontexten (vgl. zum Beispiel [Rei91]). Eine formelle Beschreibung und Definition der Koreferenz soll in dieser Arbeit nicht gegeben werden.

1.3 Aufbau der Arbeit

Die in dieser Arbeit vorgestellten Ansätze zur Integration von Wissen aus der Perzeption in einen aus einer Instruktion stammenden Wissensbestand sollen im Geometrischen Agenten zum Einsatz kommen. Aus diesem Grund wird in Kapitel 2 zunächst das Modell des Geometrischen Agenten vorgestellt und in bekannte Agentenmodelle eingeordnet. In Kapitel 3 werden dann die Modelle der Wissensrepräsentation vorgestellt, die dem Geometrischen Agenten als kognitive Karten dienen und die die Grundlage für die folgenden Kapitel darstellen. In Kapitel 4 wird dann ein Ansatz vorgestellt, der es ermöglicht ein Maß der Ähnlichkeit von den in Abschnitt 3.3 vorgestellten Graphen zu bestimmen. In Kapitel 6 wird ein Vorschlag gemacht, wie dieses Maß der Ähnlichkeit zur Bestimmung der Koreferenz von Knoten während der Navigation des Geometrischen Agenten verwendet werden kann. Da es zum Verständnis von Kapitel 6 nötig ist, die in den Graphen verwendeten räumlichen Konzepte zu kennen, wird zuvor in Kapitel 5 dargestellt, wie sich räumliches Wissen mit den in Abschnitt 3.3 vorgestellten Graphen repräsentieren lässt und welche Konzepte benötigt werden.

Kapitel 2

Der Geometrische Agent

Wie bereits in der Einleitung erwähnt, dienen die in dieser Arbeit vorgestellten Prozesse zur Integration von aus der Perzeption stammenden Wissen in einen instruierten Wissensbestand im Geometrischen Agenten (GA). Der GA dient als ein formales Rahmenwerk zur Untersuchung von Zusammenhängen zwischen der räumlichen Information einer im Voraus gegebenen Wegebeschreibung und der, aus der Perzeption eines Agenten gewonnenen, räumlichen Information.

Er soll als solches eine Wegbeschreibung von einem Benutzer erhalten und in der Lage sein, aufgrund dieser selbstständig in seiner Umgebung zu navigieren. Der GA simuliert damit einen künstlichen Agenten wie zum Beispiel einen Roboter. Für die Simulation ist der Ablauf in zwei grundlegende Phasen unterteilt: Die Instruktionsphase und die Aktionsphase. Im Folgenden soll zunächst ein Überblick über das Feld der Agentensysteme nach Russell und Norvig ([RuN03]) gegeben werden. Danach wird die Architektur des GA beschrieben und in das Feld der Agentensysteme eingeordnet.

2.1 Agentensysteme

Das Konzept des Agenten wird im Bereich der künstlichen Intelligenz häufig als von zentraler Bedeutung angesehen. Ein Agent wird zum Beispiel von Russell und Norvig beschrieben als alles, was seine Umgebung durch Sensoren wahrnimmt und mit seiner Umgebung interagiert (vgl. Abbildung 2.1 und [RuN03]). Die Umgebung eines Agenten kann verschiedener Art sein. So kann die Umgebung eines Softwareagenten das World Wide Web sein und die Umgebung eines Roboters die reale Welt.

Aktuatoren sind in diesem Zusammenhang alles womit der Agent Aktionen in seiner Welt ausführen kann. Die Aktuatoren eines Roboters können zum Beispiel ein Fahrgestell für die Fortbewegung oder Greifarme für die Manipulation von Objekten in seiner Umgebung sein. Die Perzeption sei im Folgenden, der durch die Sensoren wahrgenommene Input eines Agenten zu

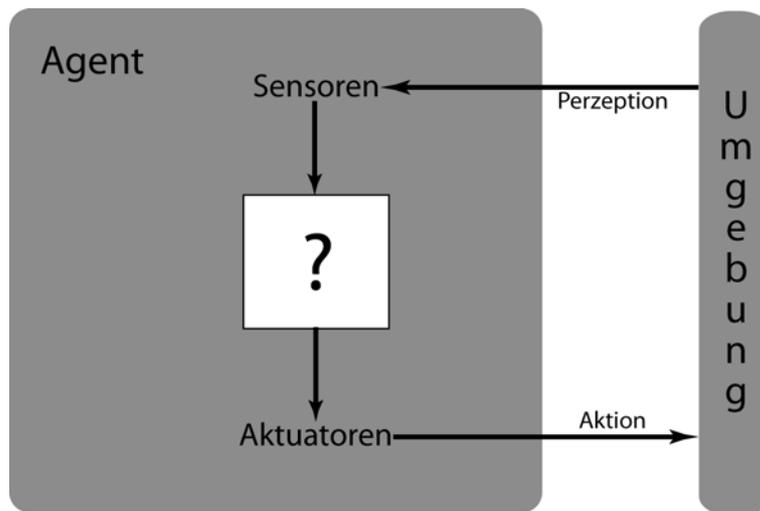


Abbildung 2.1: Basismodell eines Agenten nach [RuN03], der durch Sensoren und Aktuatoren mit der Umwelt interagiert.

einem bestimmten Zeitpunkt, eine **Perzeptionssequenz** alles, was der Agent bis zu einem bestimmten Zeitpunkt wahrgenommen hat. Die Verbindung zwischen Perzeption und Aktion stellt nach [RuN03] die so genannte **Agentenfunktion** dar. Die Agentenfunktion bildet von der Perzeption oder einer Perzeptionssequenz auf eine oder mehrere Aktionen ab. Betrachtet man den Agenten als kognitives Modell, modelliert die Agentenfunktion das Gehirn des Agenten.

In der Literatur wird häufig von einem rationalen oder vernünftig handelnden Agenten gesprochen, also von einem Agenten, dessen Handlungen einem Menschen vernünftig erscheinen. Die Abbildung von der Perzeption auf die Handlungen macht keinerlei Aussagen über die Rationalität der Handlungen. Die ausgeführte Handlung kann folglich unvernünftig erscheinen. Es muss also die Möglichkeit geben, möglichst objektiv⁵ zu beurteilen, ob eine Handlung vernünftig ist. Aufgrund dieser Beurteilung kann dann auch die Definition einer Agentenfunktion erfolgen, die auf vernünftige Handlungen abbildet.

Russell und Norvig beschreiben in [RuN03] einen rationalen Agenten zuerst einmal sehr allgemein als einen Agenten, der 'das Richtige' tut. Diese ungenaue Beschreibung macht nur dann Sinn, wenn beurteilt werden kann, was 'das Richtige' ist. Intuitiv lässt sich sagen, dass 'das Richtige' eine Handlung ist, die erfolgreich bezüglich der Ziele des Agenten ist. Um eine Handlung als erfolgreich einstufen zu können, muss also das Ziel des Agenten bekannt sein. Sind die Ziele bekannt, so können anhand der Ziele Kriterien bestimmt werden, die zur Beurteilung des Erfolgs verwendet werden können. Diese Kriterien werden von Russell und Norvig als das **Performanzmaß** beschrieben. Damit lässt sich die in [RuN03] auf S. 36 gegebene Definition eines rationalen Agenten formulieren.

⁵ Genau genommen muss es inter-subjektiv heißen.

„For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.” ([RuN03] S. 36)

Wichtig ist zu beachten, dass Rationalität keinesfalls bedeutet, dass der Agent immer die objektiv richtige Entscheidung trifft oder durch seine Handlungen letztlich seine Aufgabe erfüllt. Die Entscheidung muss nur bezüglich seiner momentanen Perzeption und seines momentanen Wissens das Performanzmaß maximieren. Die Entscheidung sollte uns (als Mensch) unter diesen Kriterien vernünftig erscheinen.⁶

Ein Agent muss auf einem physischen System, wie zum Beispiel einem Roboter oder einem Rechner (bei einem Softwareagenten), implementiert sein. Die Implementation setzt ein von Russel und Norvig so genanntes **Programm** oder allgemeiner gesagt einen Algorithmus voraus. Das gewählte Programm für den Agenten muss zu dem gewählten System passen. So darf zum Beispiel ein Programm keine Handlungen verlangen, die das System nicht ausführen kann. Ist der Agent ein Roboter mit Rädern, so ist es unsinnig, wenn das Programm die Aktion *hinsetzen* fordert. Dies mag im ersten Moment trivial und nicht erwähnenswert erscheinen, ist aber insbesondere dann wichtig, wenn eine domänenunabhängige und austauschbare Repräsentation von Wissen verwendet wird. In einer solchen Repräsentation ist Wissen repräsentiert ohne Rücksicht auf die Architektur und die Umgebung des sie verwendenden Systems, wodurch es durchaus zu dem eben genannten Problem kommen kann.

Eine solche domänenunabhängige Repräsentation von Wissen ist keinesfalls selbstverständlich, jedoch wünschenswert. In den meisten der in der Praxis verwendeten Agenten ist das Wissen implizit in der Agentenfunktion enthalten, also in dem Teil des Agenten, der für die Ausführung der Aktionen aufgrund der momentanen Perzeption zuständig ist. So hat der Autopilot eines Flugzeugs kein explizit repräsentiertes Wissen über Flugzeuge, deren Steuerelemente, Instrumente oder die Physik des Fliegens, das er dann verwendet, um daraus zu schließen, welche Aktionen, also Bewegungen der Steuerflächen, er auszuführen hat, um das Flugzeug eine Kurve fliegen zu lassen oder sicher zu landen. Vielmehr ist dieses Wissen von den Entwicklern des Systems implizit in das Programm eingeflossen. Da das Wissen aber nicht explizit repräsentiert ist, ist es nicht möglich, dieses Wissen für andere Zwecke zu verwenden. Dieses Vorgehen ist in höchstem Maße domänenabhängig.

Der Gegensatz zu dieser so genannten prozeduralen Darstellung ist die deklarative Darstellung. Bei der deklarativen Darstellung werden Wissen und Verarbeitung voneinander getrennt. In solchen Systemen gibt es symbolische Repräsentationen des aktuellen Problems und des Wissens des Agenten, sowie

⁶ Eine Möglichkeit eine Entscheidung als Vernünftig anzusehen ist, sie mit der Entscheidung eines Menschen zu vergleichen, der in der gleichen Situation ist. In der gleichen Situation bedeutet hierbei, dass er das gleiche (Vor-)Wissen und die gleichen Informationen über seine Umgebung hat.

ein oder mehrere möglichst allgemeine Problemlösungsverfahren, die mit der Wissensrepräsentation arbeiten. Dadurch ist es einerseits möglich, die Problemlösungsverfahren auch auf Wissen anderer Domänen anzuwenden, andererseits ist es ebenso möglich, das repräsentierte Wissen für andere Probleme zu verwenden. Agenten, die solch eine symbolische Repräsentation von Wissen verwenden, heißen wissensbasierte Agenten⁷. Die zentrale Komponente der wissensbasierten Agenten ist die Wissensbasis (WB), welche die Repräsentation des Wissens enthält.

Für die Wissensrepräsentation ist der Begriff Wissen von zentraler Bedeutung, weshalb ich an dieser Stelle eine kurze Beschreibung des Begriffs geben möchte.⁸ Wissen wird im Allgemeinen als der Inbegriff von Kenntnis definiert. Im Gegensatz dazu gibt es den Glauben oder die Überzeugung. Dem folgend unterscheidet zum Beispiel Reimers in [Rei91] zwischen dem Wissen und der Überzeugung eines Wissensträgers. Nach Reimers ist das Wissen die Menge aller vom Wissensträger als wahr angenommenen Aussagen, die auch tatsächlich wahr sind. Zur Überzeugung eines Wissensträgers gehören alle Aussagen, von denen er glaubt, dass sie wahr sind. Ich werde im Folgenden, wenn ich von Wissen spreche, sowohl das Wissen, als auch die Überzeugung eines Wissensträgers meinen.

Nach [BrL85] ist Wissensrepräsentation das Aufschreiben von Symbolen, die in einer erkennbaren Weise einem Ausschnitt einer zu repräsentierenden Welt entsprechen. Was 'erkennbar' bedeutet, wird von Brachmann und Levesque nicht näher spezifiziert. Reimers erweitert in [Rei91] deshalb den Begriff und verlangt, „daß das Aufschreiben von Strukturen erst dann eine (Wissens-) Repräsentation ist, wenn es eine Interpretationsvorschrift gibt, die mindestens die Formulierung und Auswertung von Anfragen auf diesen Strukturen zuläßt.“ ([Rei91] S. 9)

Dies führt zu folgender Definition von Wissensrepräsentation, in der der Begriff Interpretationsvorschrift durch den m. E. treffenderen Ausdruck Repräsentationsabbildung ersetzt wurde:

„Eine (Wissens-)Repräsentation (WR) ist eine Menge von Zeichen, die mittels einer Repräsentationsabbildung auf die Merkmale des repräsentierten Weltausschnitts (Repräsentandum) abgebildet werden.“ ([Rei91] S. 10)

Habel spricht in [Hab86] im Zusammenhang des repräsentierten Weltausschnitts (Domäne) von einer projizierten Welt (vgl. Abbildung 2.2), was zum Ausdruck bringt, dass die Wissensrepräsentation nicht die reale Welt repräsentiert, sondern das Wissen des Systems über die Welt, welches auch

⁷ Man spricht heute nicht mehr von Systemen mit künstlicher Intelligenz, sondern von wissensbasierten Systemen, da die Systeme zwar Wissen(sbasen) verwenden, aber von echter Intelligenz weit entfernt sind (vgl. z. B. [Rei91]).

⁸ Eine hinsichtlich philosophischer und psychologischer Gesichtspunkte adäquate Definition von Wissen kann an dieser Stelle selbstverständlich nicht gegeben werden.

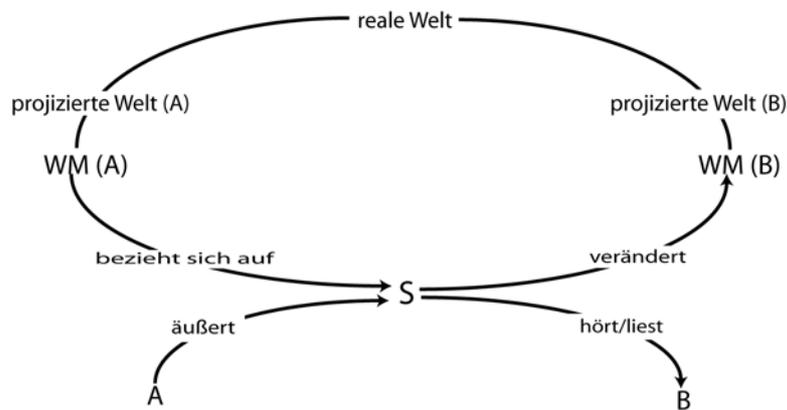


Abbildung 2.2: Eine Kommunikationssituation nach [Hab86] mit einem von A geäußerten Satz S.

Aussagen enthalten kann, die nicht der realen Welt, sondern nur der Überzeugung des Systems entsprechen. Die während der Instruktionsphase des GA erzeugte Wissensrepräsentation ist vollkommen aus kommuniziertem Wissen gebildet. Da Wissen, welches zwischen zwei Wissensträgern kommuniziert wird, nur der projizierten Welt des Senders A entspricht und nicht klar ist, wie dieses entstanden ist bzw. ob die Kommunikation fehlerfrei war, hat dieses Wissen den Charakter, besonders unsicher zu sein (vgl. Abbildung 2.2).

2.2 Instruktionsphase

Während der Instruktionsphase erhält der Agent die Wegbeschreibung, mit deren Hilfe er navigieren soll. Generell führt Kommunikation zwischen Menschen oder im Idealfall auch zwischen wissensbasierten Systemen zu einer Veränderung im Wissensbestand des Wissensträgers. Ein Sprecher A, der einen Satz S an B richtet, bezieht sich dabei auf seine WB und verändert die WB von B, wobei beide Wissensbasen die projizierten Welten der Wissensträger repräsentieren. Die aus [Hab86] entnommene Abbildung 2.2 soll diesen Vorgang verdeutlichen.

Während der Instruktionsphase des GA wird die Wegbeschreibung interpretiert und in eine Repräsentation transformiert, die dann Teil der WB des Agenten⁹ wird. Diese Repräsentation heißt **Instruktionsmodell** (vgl. Abbildung 2.3). Die Interpretation der Wegbeschreibung basiert auf linguistischem Wissen sowie Wissen über zeitliche und räumliche Konzepte. Dieses Wissen stellt einen Teil der Repräsentationsabbildung dar (vgl. Abschnitt 2.1). Der GA führt die Transformation in zwei Schritten aus.

⁹ Ich spreche hier von einem Teil der WB, da der GA bereits eine WB in Form der GCS und des semantischen und syntaktischen Wissens hat (vgl. Abb. 2.3). Er soll im Vorwege aber kein Wissen über die Route besitzen.

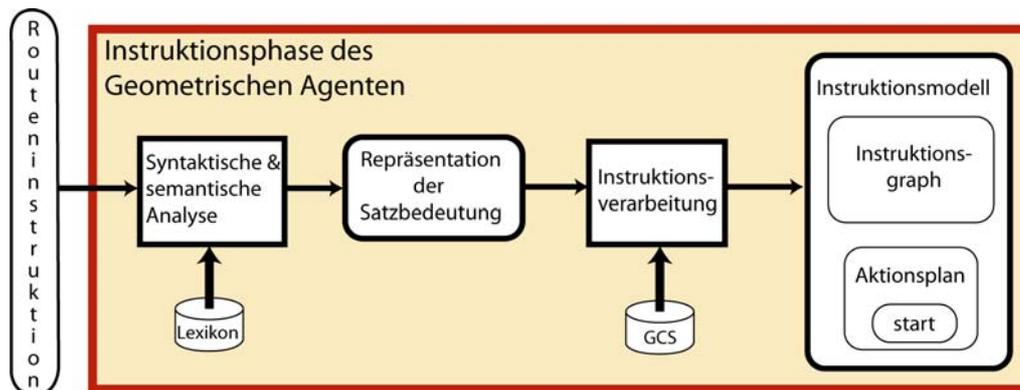


Abbildung 2.3: Aufgaben während der Instruktionsphase des Geometrischen Agenten. (Entnommen aus [TSH02].)

Im ersten Schritt wird die Wegbeschreibung syntaktisch und semantisch analysiert und in eine konzeptuelle Sprache übersetzt. Das dabei verwendete Lexikon ist die linguistische WB. Das Lexikon enthält linguistisches Wissen und Wissen über zeitliche und räumliche Konzepte. Mit Hilfe des Lexikons werden zum Beispiel Verben kategorisiert und diese Kategorien einheitlich repräsentiert. In Abschnitt 5.2 wird noch genauer auf diesen Vorgang eingegangen.

Im zweiten Schritt transformiert die Instruktionsverarbeitung die konzeptuelle Sprache in zwei Repräsentationen, welche zusammen die Conceptual Route Instruction Language (CRIL) darstellen, zum einen in einen CRIL-Graphen, der das räumliche Wissen repräsentiert (Instruktionsgraph), und zum anderen in einen so genannten Aktionsplan, der imperative Ausdrücke, also Anweisungen etwas zu tun, in einer konzeptuellen Sprache enthält. Durch den Aktionsplan werden die temporalen Aspekte der Wegbeschreibung repräsentiert, also in welcher Reihenfolge welche Aktionen ausgeführt werden sollen.

Während der Transformation in den CRIL-Graphen wird von der so genannten Geometric Concept Specifications (GCS) Gebrauch gemacht. Sie enthält Wissen über die Geometrie von räumlichen Konzepten der natürlichen Sprache wie es z. B. in [ETH00] von Eschenbach et al. spezifiziert wird. Dabei wird die Repräsentation der räumlichen Konzepte in eine äquivalente Repräsentation in geometrischen Begriffen transformiert. Folgendes Beispiel soll dies verdeutlichen.

Die räumliche Präposition *zu*, wie sie zum Beispiel in dem Satz „der Weg führt zu dem roten Haus“ verwendet wird, impliziert einen Pfad¹⁰, dessen Endpunkt in der Nähe eines roten Hauses lokalisiert ist, während sein Startpunkt grade nicht in der Nähe des roten Hauses lokalisiert ist. Die Repräsentation der Präposition *zu* in der Repräsentation der Satzbedeutung wird so mittels der GCS in die Repräsentation eines Pfades mit einem Start- und einem Endpunkt

¹⁰ Das Konzept Pfad sei hier informell verwendet. Eine genauere Spezifikation folgt in Abschnitt 5.2.

transformiert, wobei die Lokalisationen der Punkte so repräsentiert werden, dass sie den Vorschriften der GCS entsprechen.

Zusätzlich zu den eben beschriebenen Transformationen können während der Instruktionsverarbeitung bestimmte Inferenzen vorgenommen werden. So sind zum Beispiel Start- und Endpunkte von verschiedenen Pfaden nach der ersten Transformation immer durch unterschiedliche Punkte repräsentiert. Die Reihenfolge der Pfade ist durch die Anweisungen im Aktionsplan repräsentiert. Es kann aber sinnvoll sein, End- und Startpunkt von aufeinander folgenden Pfaden zu einem Punkt zu vereinen. Auf diesen Punkt wird in Abschnitt 5.4 noch eingegangen.

2.3 Die Umgebung des Agenten

Aufgabe des Agenten ist es, Informationen der Instruktion mit aus der Umwelt gewonnenen Informationen zu vergleichen und daraufhin in seiner Umwelt zu navigieren. Diese Aufgabe ist als solches mit der eines mobilen Roboters vergleichbar. Der Geometrische Agent kann dies aber in einer idealisierten Weise tun, da Aufgaben wie Wahrnehmung oder Fortbewegung vom Agenten simuliert werden. Somit ist es möglich, den Realismusgrad der Umgebung zu variieren.

Um die bei der Navigation aufgrund einer Wegbeschreibung auftretenden Probleme und Anforderungen an die kognitiven Systeme adäquat untersuchen zu können, ist es nötig, die Umgebung des Agenten so zu modellieren, dass die während der Navigation auftretenden Probleme eines Agenten in einer realen Umgebung, auch in der Umgebung des Geometrischen Agenten auftreten. Beispiele für solche Probleme sind die (teilweise) Verdeckung von Objekten durch andere (höhere, aber auch niedrigere) Objekte, die Abhängigkeit der Erkennung von Objekteigenschaften¹¹ von der Entfernung eines Agenten zu einem Objekt, aber auch das eingeschränkte Blickfeld eines Agenten oder die Fortbewegung entlang eines erkannten Wegs in der Umgebung.

Der GA navigiert in einer virtuellen Umgebung. In seiner Umgebung werden Objekte wie Häuser, Bäume oder Wege durch geometrische Objekte, wie zum Beispiel Polygone, dargestellt. Durch eine derartige Modellierung der Umgebung ist es möglich, bestimmte kognitive Fähigkeiten bezüglich Wahrnehmung und Fortbewegung von Robotersystemen zu simulieren, die zum gegenwärtigen Zeitpunkt nicht zur Verfügung stehen. Ohne diese Fähigkeiten der Wahrnehmung und Fortbewegung ist eine adäquate Navigation nicht möglich. Als Beispiel sei hier die Erkennung von Objekten wie Häusern oder Bäumen in realen Umgebungen angegeben. Da mit Hilfe des Geometrischen Agenten kognitive Prozesse und Strukturen des Menschen untersucht werden sollen, sollte es ihm zumindest möglich sein, Objekte

¹¹ Beispiele für solche Objekteigenschaften sind Namen, Anzahl von Stockwerken oder die Farbe bestimmter Teile des Objekts.

ähnlich genau zu erkennen wie ein Mensch. Eine solche Genauigkeit wird, vor allem in natürlichen Umgebungen, von heutigen Bilderkennungssystemen nicht erreicht.

Da in der Simulation variiert werden kann, wie gut der Agent bestimmte Aufgaben lösen kann, kann der GA auch verwendet werden, um zu prüfen, in welchem Maß die kognitiven Fähigkeiten der Wahrnehmung und Fortbewegung in Robotersystemen gegeben sein müssen, um eine adäquate Navigation eines solchen Systems in einer realen Umgebung zu ermöglichen. Ein weiterer Vorteil einer solchen Modellierung ist, dass Versuche beliebig unter absolut identischen Bedingungen der Umgebung reproduzierbar sind. Somit können beispielsweise die Auswirkung von Veränderungen in der Interpretation der Wegbeschreibung oder Veränderungen der WR auf die Navigationsleistung untersucht werden.

Russel und Norvig stellen in [RuN03] folgende Klassifikation bezüglich der Aufgaben eines Agenten vor:

- Wenn die Handlung eines Agenten nicht von seinen früheren Handlungen beeinflusst ist, heißt das Aufgabenumfeld **episodisch**, sonst **sequentiell**.
- Wenn die für die Lösung der Aufgabe relevanten Eigenschaften der Umgebung über die Sensorik des Agenten zu jedem Zeitpunkt wahrnehmbar sind, so heißt die Umgebung **vollbeobachtbar**, sonst **teilbeobachtbar**.
- Wenn der Folgezustand der Umgebung vollständig durch den momentanen Zustand und die Handlung des Agenten bestimmt ist, so heißt die Umgebung **deterministisch**, sonst **stochastisch**.
- Wenn sich die Umgebung ändert, während der Agent überlegt, so heißt die Umgebung **dynamisch** bezüglich des Agenten, sonst **statisch** bezüglich des Agenten.
- Der Zustand der Umgebung, die Art wie Zeit gehandhabt wird und Perzeption und Aktion lassen sich als **diskret** oder **stetig** kategorisieren. So ist ein Schachspiel beispielsweise eine diskrete Umgebung, da es eine endliche Menge von eindeutigen Zuständen hat.

Das Aufgabenumfeld eines Menschen in der realen Welt ist, wie leicht zu erkennen ist, sequentiell, teilbeobachtbar, stochastisch, dynamisch und stetig. Ein solches Aufgabenumfeld hat die größte anzunehmende Komplexität. Das Aufgabenumfeld des Geometrischen Agenten unterscheidet sich darin, dass die Umgebung deterministisch und statisch ist, da keinerlei Veränderungen der Umgebung außer der Bewegung des Agenten vorgesehen sind.

Die Teilbeobachtbarkeit der Umgebung des Geometrischen Agenten ergibt sich direkt aus dem eingeschränkten Sichtfeld des Geometrischen Agenten sowie der Verdeckung von Objekten in der Umgebung. Dass das Aufgabenumfeld sequentiell ist, ist nicht direkt erkennbar. So wäre es bei einer vollständig korrekten Wegbeschreibung und einer vollbeobachtbaren Umgebung möglich, die Handlung des Agenten nur von der momentanen

Wahrnehmung und der internen Repräsentation der Wegbeschreibung abhängig zu machen. Dies wäre möglich, da unter den gegebenen Voraussetzungen garantiert werden könnte, dass jeder Teil der Wegbeschreibung eindeutig in der Umgebung identifiziert werden könnte und somit jede Entscheidung während der Navigation korrekt getroffen werden könnte. Zum einen ist eine natürlichsprachliche Wegbeschreibung nicht zwangsläufig korrekt und zum anderen ist die Umgebung eben nicht vollbeobachtbar. Dies führt zu Situationen während der Navigation, in denen es dem Navigator nicht möglich ist, den als nächstes anstehenden Teil der Wegbeschreibung in der Umgebung eindeutig zu identifizieren. In einer solchen Situation soll der Agent trotzdem handlungsfähig bleiben. Es muss ihm möglich sein, eine Entscheidung zu treffen, weiter zu navigieren, aber gleichzeitig im Gedächtnis zu behalten, dass es sich um eine unsichere Entscheidung handelte. Die folgenden Entscheidungen müssen dann unter der Berücksichtigung der vorigen unsicheren Entscheidung getroffen werden. Diese mögliche Abhängigkeit einer Entscheidung von vorherigen Entscheidungen entspricht einem sequentiellen Aufgabenumfeld.

Die Umgebung des Geometrischen Agenten ist genau genommen diskret, da sie auf einem Rechner implementiert ist. Sie repräsentiert aber einen stetigen Raum und Repräsentation und Berechnung verwenden stetige Modelle. Folglich muss die Umgebung als stetig angesehen werden (vgl. zum Beispiel [RuN03] S. 42).

Die Einordnung der Sensorik und Aktorik des Geometrischen Agenten ist nicht so einfach möglich. Sensorik und Aktorik des Geometrischen Agenten abstrahieren von metrischen Konzepten. So sind die Konzepte, die das Programm des Geometrischen Agenten von der Sensorik erhält und für die Aktorik verwendet, diskrete und qualitative Konzepte¹². Es ist dem GA zum Beispiel nicht möglich, die genaue Entfernung zweier Objekte wahrzunehmen oder zu erfassen, in welchem Winkel zwei Wege zueinander stehen. Ebenso ist es dem GA nicht möglich, eine Anweisung wie „Fahre 10,2 Meter geradeaus und wende dann 23° entgegen dem Uhrzeigersinn“ auszuführen. Er ist aber in der Lage zu erkennen, ob ein Objekt in einer Anordnungsrelation wie *links* oder einer topologischen Relation wie *in* zu einem anderen Objekt steht sowie zum Beispiel einem Weg zu folgen. Dies könnte zu der Einschätzung führen, dass Sensorik und Aktorik des Geometrischen Agenten diskret sind. Da Perzeption und Aktorik aber in einer stetigen Umgebung stattfinden und auf einer dem 'Bewusstsein' des GA nicht zugänglichen Ebene ebenfalls stetige Berechnungen verwenden, sind auch Perzeption und Aktorik als stetig anzusehen.

Da der Folgezustand der Umgebung nur von dem momentanen Zustand und der nächsten Aktion des Geometrischen Agenten abhängt, ist die Umgebung nach [RuN03] deterministisch. Außerdem ist sie statisch, da sie sich nicht verändert, wenn der Agent sich nicht bewegt. Nun ließe sich argumentieren, dass eine solche Umgebung nicht den nötigen Realitätsgrad aufweist, um

¹² Vgl. diesbezüglich [Hel02].



Abbildung 2.4: Karte des FB Informatik der Universität Hamburg als Beispielumgebung des GA (rot gekennzeichnet) und seinem Blickfeld (grau schattiert).

daraus Schlüsse bezüglich der Navigation eines Agenten in einer realen Umgebung ziehen zu können. Allerdings sind die meisten für die Navigation verwendeten Objekte der Umgebung auch in der realen Welt statisch. Dies gilt sowohl für Objekte entlang derer ein Agent sich bewegt, als auch für Objekte, die als Navigationshilfen, also für die Identifizierung von bestimmten Orten verwendet werden. Fahrende Autos zum Beispiel stellen während der Bewegung in einer Umgebung sicherlich erhöhte Anforderungen an den Agenten, sind für die eigentliche Navigation, also das Finden des richtigen Wegs, aber nicht von Interesse. Die dynamischen, stochastischen Aspekte der realen Welt scheinen also für die Navigationsleistung nicht von Interesse.

2.4 Navigationsphase

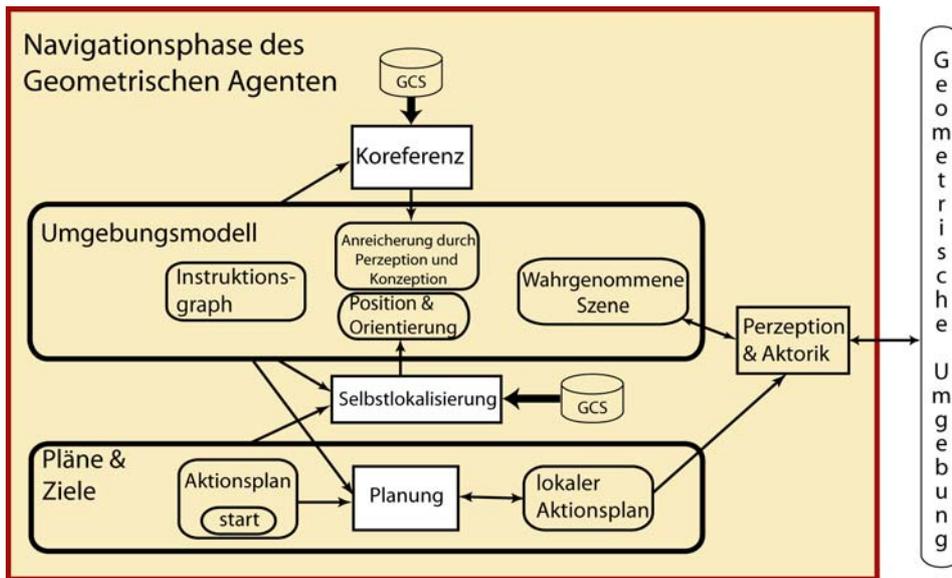
Abbildung 2.5 aus [TSH03] verdeutlicht die Architektur des GA in der Navigationsphase. Das für Perzeption und Aktorik verwendete Modul wurde im vorangegangenen Abschnitt bereits erläutert. Um nun in der Umgebung navigieren zu können, muss der Agent das Instruktionsmodell (vgl. Abschnitt 2.2) mit dem aus der Wahrnehmung erstellten Modell (Perzeptionsmodell)

vergleichen und zwischen den Knoten der beiden Graphen eine Verbindung herstellen, die das gleiche Objekt der Umgebung zu repräsentieren scheinen. Diese Aufgabe übernimmt das **Koreferenzmodul**, in dem die in dieser Arbeit vorgestellten Prozesse verwendet werden. Da das Koreferenzmodul die Perzeption steuern können muss, hat auch das Koreferenzmodul Zugriff auf das Wahrnehmungsmodul.¹³ Ist eine Identifizierung von Knoten aus Instruktions- und Wahrnehmungsmodell möglich, so wird der CRIL-Graph, also das Instruktionsmodell, mit den entsprechenden Teilen des Perzeptionsmodells angereichert. Die so entstehende WB heißt **Umgebungsmodell**.

In dem Umgebungsmodell muss die Position des Agenten repräsentiert sein. Für die korrekte Repräsentation der Position ist das **Selbstlokalisierungsmodul** zuständig. Auch dieses Modul greift auf das Modul für Wahrnehmung zu. Schließlich wird noch ein Modul gebraucht, welches den lokalen Aktionsplan dynamisch erstellt. Dies übernimmt das **Planungsmodul**. Das Planungsmodul gibt atomare Anweisungen, wie zum Beispiel, dass der Agent einem Weg folgen oder sich drehen soll.

Dies ist nötig, da verbale Wegbeschreibungen in vielen Fällen unterbestimmt sind und der nach der Instruktionsphase aufgestellte Aktionsplan nur sehr grob die auszuführenden Aktionen enthält. So enthält eine Wegbeschreibung zum Beispiel meist nur wichtige Landmarken und die Anweisung, einer Straße zu folgen, bedeutet häufig auf ihr zu bleiben, selbst wenn zwischendurch Abzweigungen vorhanden sind. Die Anweisungen im Aktionsplan enthalten dementsprechend nur die in der Wegbeschreibung enthaltenen groben Anweisungen. Das Planungsmodul entscheidet in der jeweiligen Situation, welche elementare Aktion als nächstes ausgeführt werden muss und weist dann das für Perzeption und Aktion zuständige Modul an, diese Aktion auszuführen.

¹³ Warum diese Steuerung nötig ist, wird in Kapitel 6 noch deutlich werden.



Kapitel 3

Grundlagen der Wissensrepräsentation

In diesem Kapitel sollen die formalen Grundlagen für die später vorgestellten Vergleichs- und Integrationsprozesse vorgestellt werden. Zunächst wird im folgenden Abschnitt eine kurze Einführung in die Theorie der referentiellen Netze gegeben. Referentielle Netze sind die Grundlage der CRIL und werden in der Instruktionsphase des GA verwendet. Der folgende Abschnitt soll ebenfalls als Einführung in die Wissensrepräsentation mit logikbasierten Sprachen dienen. In Abschnitt 3.2 werden Taxonomien formal eingeführt. Sie dienen der Bestimmung der Ähnlichkeit von Konzepten und werden deshalb für die in Abschnitt 3.3 vorgestellten CRIL-Graphen benötigt.

3.1 Referentielle Netze

Die Conceptual Route Instruction Language wurde entworfen auf Grundlage der in [Hab88] vorgestellten Erweiterung der referentiellen Netze für die Darstellung von räumlichem Wissen. Referentielle Netze wurden ursprünglich für die Textproduktion und Textanalyse¹⁴ entwickelt und werden zu diesem Zweck auch in der Instruktionsphase des GA verwendet (vgl. Abschnitt 2.2). Grundsätzlich sollte es so möglich sein, CRIL-Graphen zusammen mit einer aus SRL-Ausdrücken bestehenden Wissensbasis zu verwenden. So lassen sich zum Beispiel die im folgenden Abschnitt eingeführten domänenunabhängigen Regeln auch für CRIL-Graphen verwenden. Im Folgenden werde ich deshalb kurz die Theorie der referentiellen Netze darstellen.

¹⁴ Für eine ausführliche Einführung in die Theorie der referentiellen Netz vgl. [Hab86],

3.1.1 Die Repräsentationssprache SRL

Grundlage der referentiellen Netze ist die Semantic Representation Language (SRL). Sie ist eine propositionale, logik-orientierte Repräsentationssprache, in der zwischen Termen und Formeln im Sinne der Logik unterschieden wird. Die SRL wurde ursprünglich als Bedeutungsrepräsentationssprache für die natürlichsprachliche Fragebeantwortung konzipiert und später zu einer Repräsentationssprache für Textbedeutung erweitert. Um für diesen Zweck als adäquater Repräsentationsformalismus zu dienen, muss die SRL den Ansprüchen einer (allgemeinen) Wissensrepräsentationssprache genügen, worauf bei ihrer Entwicklung besonders Wert gelegt wurde¹⁵.

Im Folgenden seien nun zunächst die syntaktischen Grundlagen der SRL beschrieben. Das Basisinventar der SRL entspricht in groben Zügen dem bekannten Inventar der Prädikatenlogik erster Stufe. Habel unterscheidet allerdings nicht explizit zwischen Prädikaten und Funktionen, sondern fasst beide zu dem Konzept der Operatoren zusammen. Die Unterscheidung erfolgt dann über den so genannten Effekt eines Operators, der jedem Operator zugeordnet ist und der zum Ausdruck bringt ob der Operator term- oder formelbildend ist. Konstanten sind in der SRL grundsätzlich Namen und durch Hochkommas gekennzeichnet.

Definition 3.1.1: *Die syntaktische Symbole (das Alphabet) der SRL besteht aus den Mengen der verfügbaren Symbole*

- *VAR, einer abzählbaren Menge der Termvariablen,*
- *OP, einer endliche Menge der Operatoren und*
- *EFF := {t,f} einer Menge der Effekte von Operatoren, wobei t für Terme und f für Formeln steht,*

sowie den logischen Symbolen der

- *Junktoren: $\neg, \wedge, \vee, \Leftrightarrow, \Rightarrow$*
- *Quantoren: \exists, \forall ,*
- *Hilfssymbolen: $), (, ,$*

Die Formationseigenschaft der Operatoren wird durch die Typ-Funktion spezifiziert:

$$\text{TYP: OP} \rightarrow \text{EFF} \times \mathbb{N} \times \mathbb{N} \times \mathbb{N}$$

¹⁵ Bezüglich der Ansprüche an eine Wissensrepräsentationssprache vgl. z.B. [Hab86] Abschnitt 2.2.

Das durch $TYP(op)$ erzeugte Quadrupel sei im Folgenden immer durch $TYP(op) = \langle eff, m, n, p \rangle$ bezeichnet. Hierbei ist eff der durch den Operator erzeugte Effekt, also ist eff gleich t , wenn es sich um einen termbildenden oder gleich f , wenn es sich um einen formelbildenden Operator handelt (vgl. oben). Die Variable m gibt an wie viele Variablen durch op gebunden werden, n besagt wie viele Terme und p wie viele Formeln op als Argument hat. Damit lassen sich die Mengen TER der Terme und FOR der Formeln rekursiv nach folgendem in [Hab86] gegebenen Definitionsschema definieren.

Definitionsschema 3.1.2: Die Menge der Terme TER und die Menge der Formeln FOR lässt sich über folgendes induktives Schema definieren:

- Jede Variable ist ein Term.
- Wenn $op \in OP$ und $typ(op) = \langle t, m, n, p \rangle$ mit
 - $x_1, \dots, x_m \in VAR$
 - $t_1, \dots, t_n \in TER$
 - $f_1, \dots, f_p \in FOR$
 , dann ist $op(x_1, \dots, x_m, t_1, \dots, t_n, f_1, \dots, f_p) \in TER$.
- Dies sind alle Terme.
- Wenn $op \in OP$ und $Typ(op) = \langle f, m, n, p \rangle$ mit
 - $x_1, \dots, x_m \in VAR$
 - $t_1, \dots, t_n \in TER$
 - $f_1, \dots, f_p \in FOR$
 , dann ist $op(x_1, \dots, x_m, t_1, \dots, t_n, f_1, \dots, f_p) \in FOR$.
- Dies sind alle Formeln.

Eine Definition sei hier nicht gegeben. Sie entspräche aber den bekannten Definitionen der Semantik der Prädikatenlogik, wie sie zum Beispiel in [Sch95] oder [Ben01] gegeben sind. Zu beachten ist lediglich, dass es anstelle von Funktionen und Prädikaten (vgl. oben) termbildende und formelbildende Operatoren gibt.

In der SRL werden sowohl Terme als auch Formeln als wohlgeformte (SRL-) Ausdrücke bezeichnet. Um eine einfache Unterscheidung zwischen Termen und Formeln zu ermöglichen, wird die SRL als Paar von Mengen aufgefasst:

$$SRL = \langle TER, FOR \rangle$$

Habel sagt nun bezüglich der Frage, was durch SRL-Ausdrücke repräsentiert werden soll, folgendes aus.

- SRL-Ausdrücke sollen mentale Zustände und (hierüber oder direkt) die reale Welt¹⁶ repräsentieren.
- SRL-Operatoren sollen Konzepte repräsentieren.

¹⁶ In einem Agent ist die reale Welt die Umgebung des Agenten.

Was ein mentaler Zustand genau ist, wird in [Hab86] nicht und soll auch hier nicht geklärt werden. Habel verweist auf Arbeiten aus den psychologischen und philosophischen Bereichen, insbesondere auf eine Arbeit von Pylyshyn (vgl. [Pyl80]), der auf den repräsentationellen Charakter von mentalen Zuständen hinweist. Der Zusammenhang zwischen einer Wissensrepräsentation und der realen bzw. projizierten Welt wurde bereits in Abschnitt 2.1 erläutert.

Der Begriff **Konzept** wird von Habel nicht konkret definiert. Vielmehr erläutert Habel wesentliche Aspekte des Begriffs. Ein Konzept betrifft nach [Hab86] die Gesamtheit der Bedeutungen, einschließlich der korrespondierenden Erwartungen und Standardannahmen von Begriffen. Als Bedeutung ist hier eher die kognitive und psychologische Bedeutung gemeint als die der logischen Semantik. Zusammenfassend erläutert Habel den Begriff folgendermaßen:

„Konzepte betreffen wohl strukturierte Gesamtheiten von Wissen und beinhalten insbesondere Beziehungen zu anderen Konzepten. Außerdem sind Konzepte als kognitive Entitäten anzusehen, die auf individuelle informationsverarbeitende Systeme untersucht werden. ([Hab86] S. 64)“

Die Aussage, dass SRL-Operatoren Konzepte repräsentieren (vgl. oben), bedeutet, dass auch Relationen, also mehrstellige Operatoren, Konzepte repräsentieren. Reimers definiert in [Rei91] ein Konzept als ein 3-Tupel, bestehend aus dem Konzeptnamen, der Extension und der Intension. Die Extension ist die Menge aller Objekte, die zu einem Konzept gehören. Die Intension bestimmt, welche Merkmale ein Objekt aufweisen muss, um zu einem Konzept zu gehören. Zu den Merkmalen gehören Beziehungen zu anderen Konzepten sowie Eigenschaften eines Objekts. Geht man davon aus, dass die Extension eines Konzepts eine Menge von Objekten der realen Welt ist, ist eine Relation nicht unbedingt ein Konzept. Zu beachten ist aber, dass die Menge der zu einem Konzept gehörenden Objekte nicht unbedingt (nur) Objekte der realen Welt enthalten muss. Die Menge kann auch Objekte des Denkens enthalten.

Das in [Hab86] häufig verwendete Konzept „fliegen“ soll dies verdeutlichen. Ein Satz wie „Hans fliegt nach Mallorca“ ließe sich durch die Relation *fliegen*(‘Hans’, ‘Mallorca’) repräsentieren. Allerdings lässt sich keine Menge von Objekten der realen Welt für das Konzept bestimmen, da es keine klar bestimmbar Objekte der Welt repräsentiert. Dass Fliegen ein eigenständiges Konzept ist, sollte aber intuitiv klar sein. Also liegt die Vermutung nahe, dass auch das Fliegen eines Menschen zu einem Ort im Speziellen und das Fliegen im Allgemeinen ein Objekt des Denkens ist¹⁷. Eine eingehendere

¹⁷ Beispielsweise könnte die Relation *fliegen* des Beispiels eine umfassendere Repräsentation des Konzepts *fliegen* in der WB des Wissensträgers referieren. In dieser WB könnte Fliegen als ein Objekt repräsentiert sein, das auch die Relation beschreibt.

philosophische und kognitionswissenschaftliche Betrachtung kann nicht Gegenstand dieser Arbeit sein. An dieser Stelle sei aber festgehalten, dass auch mehrstellige Prädikate Konzepte repräsentieren können.

Abschließend soll noch auf Eigenschaften von Konzepten eingegangen werden. Eigenschaften sind nach Reimers (vgl. [Rei91]) nicht näher bestimmte Angaben, über die im weiteren keine weiteren Aussagen gemacht werden sollen. Das bedeutet, dass sie nicht als eigenständiges Konzept eingeführt werden sollen, das dann wieder durch Beziehungen und eigene Eigenschaften spezifiziert werden kann. Ein Beispiel sind Farben. In vielen Domänen reicht es aus, Farben als Eigenschaftsklasse ‚Farbe‘ eines Konzepts zu repräsentieren. Die Eigenschaftsklasse lässt sich beispielsweise durch ein Prädikat $\text{farbe}(x)$ repräsentieren, dessen Argument x eine Konstante ist, die die entsprechende Farbe repräsentiert. Dann lässt sich über eine Farbe aber keine weiteren Aussagen treffen. In anderen Domänen müssen aber noch weitere Aussagen über Farben gemacht werden können. Zum Beispiel ist in der Physik wichtig, welche(r) Wellenlänge(nbereich) der Farbe zugeordnet wird.

Die obige Definition für wohlgeformte Ausdrücke lässt prinzipiell für jeden Operator jede beliebige Argumentbelegung zu. Zum Beispiel wäre $\text{links_von}(\text{kohlenstoff}, \text{haus})$ zulässig, obwohl unsinnig¹⁸. Die von Habel vorgeschlagene Möglichkeit, solche unsinnigen Belegungen zu erkennen und zu vermeiden, besteht darin, für alle Argumentstellen explizite Wohlgeformtheitsbedingungen (im folgenden Sortenbedingungen) zu formulieren.

Dies führt in [Hab86] zum Begriff der mehrsortigen Logik. Solche mehrsortigen Logiken werden in der Mathematik (und Informatik) schon seit geraumer Zeit verwendet (vgl. z.B. [Sch38]). Für mehrsortigen Logiken gilt im Allgemeinen folgendes.

- Jedem Termsymbol wird eine Sorte zugeordnet. Dabei ist sowohl eine disjunkte als auch eine nicht-disjunkte Zuordnung möglich.
- Für jeden Operator ist festzulegen, zu welchen Sorten seine Argumente gehören.
- Für jede Sorte existieren entsprechende Termvariablen, um sortenkorrekt quantifizieren zu können.

Die Sortierung wird in [Hab86] folgendermaßen vorgenommen. Das Grundinventar der SRL wird um eine endliche Menge S von Sortennamen erweitert. Den Termen der SRL wird rekursiv eine Sortenzugehörigkeit durch die Funktion $\text{sor-z}(t)$ zugeordnet.

$\text{sor-z: TER} \rightarrow S$

¹⁸ Es gibt wohl Kontexte, in denen die Belegung sinnvoll sein kann. Es steht m. E. aber außer Frage, dass sie in den meisten Zusammenhängen sinnlos ist.

Des Weiteren wird für termbildende Operatoren ein Sorten-Effekt-Operator angegeben, der mittels der Sorten der Argumente die Sorte des resultierenden Terms berechnet. Habel gibt in [Hab86] folgendes Beispiel zur Verdeutlichung an:

Wenn $\text{sor-z}(\text{Franz}') = \text{,mensch'}$,
dann $\text{sor-z}(\text{vater_von}(\text{,franz'})) = \text{mensch}$

So ergibt sich eine Zugehörigkeit von TYP zu einem so genannten Sortenzugehörigkeitsraster:

Wenn
 $\text{typ}(\text{op}) = \langle \text{eff, m, n, p} \rangle$,
dann gibt es ein Sortenraster
 $\text{arg-sorten}(\text{op}) = \langle s_1, \dots, s_n \rangle$
so dass gilt:
 $\text{op}(x_1, \dots, t_1, \dots, t_n, f_1, \dots)$ ist sortenkorrekt,
wenn $\text{sor-z}(t_i)$ verträglich mit s_i für $i=1, \dots, n$ ist.

Das bedeutet aber nicht, dass sorteninkorrekte Ausdrücke prinzipiell verboten wären. Es wird nur ein Mittel zum Erkennen solcher Inkorrektheit zur Verfügung gestellt.

Der Begriff der Verträglichkeit im obigen Raster bezieht sich auf die Verwendung einer zusätzlichen Repräsentation. Habel stellt dafür Sortenhierarchien vor. So kann beispielsweise ausgedrückt werden, dass die Sorte Mensch mit der Sorte Lebewesen verträglich ist. Damit ist ein Operator, der als Argument einen Term der Sorte Lebewesen verlangt, auch mit einem Term der Sorte Mensch wohlgeformt. Habel unterscheidet die Sortenhierarchien explizit von Is-A-Hierarchien. Is-A-Hierarchien werden hauptsächlich zur Repräsentation der Vererbung von Eigenschaften benutzt, Sorten-Hierarchien dagegen nur für die Wohlgeformtheit von Ausdrücken durch Vererbung von Sortenverträglichkeit. Auf eine Sortierung von Formeln sei hier nicht eingegangen. In Abschnitt 3.2 werden Taxonomien vorgestellt, die für die Repräsentation solcher Hierarchien verwendet werden können.

Die in [Hab86] vorgenommene Sortierung weicht insofern von den gebräuchlichen mehrsortigen Logiken (vgl. zum Beispiel [Ben01]) ab, als dass die Zuordnung der Termsymbole zu einer Sorte über eine Funktion geschieht. Üblich ist dagegen, eine Signatur zu definieren, in der die Sortenkorrektheit der Relationen und Funktion sowie die Sortenzugehörigkeit der Konstanten durch Zeichenketten repräsentiert wird, sowie für jede Sorte eine Menge von Variablensymbolen zu definieren. Durch die Unterscheidung der Variablen ist auch eine sortenkorrekte Quantifizierung möglich. Allerdings ist die Verwendung von Sortenhierarchien in diesen Ansätzen nicht üblich.

Generell sollte es in der Wissensrepräsentation möglich sein, Regeln der Form „wenn A, dann B“ oder „wenn A wahr ist, dann ist auch B wahr / dann führe B

aus“ zu repräsentieren. In der SRL kann dies durch Aussagen mit dem (Bi-) Implikationsjunktoren dargestellt werden.¹⁹ Mit Hilfe solcher Regeln lassen sich domänenabhängige Regeln, wie „wenn x ein Mensch ist, so kann x reden“ darstellen. Um auch allgemeinere Regeln, also epistemische Primitive wie zum Beispiel den Zusammenhang zwischen Transitivität und Konversität zu formulieren muss es ein höheres Sprachkonzept geben, das es ermöglicht über Operatoren zu schließen.

In [Hab86] werden dafür zusätzliche Meta-Operatoren eingeführt und somit das Sprachkonzept **Meta-SRL**. Die Meta-SRL dient der Darstellung von allgemeinen Regeln, wie zum Beispiel dem Zusammenhang zwischen Transitivität und Konversität. Es sei hier nur kurz an dem eben erwähnten Beispiel eingeführt und diene als Beispiel, wie solch allgemeingültiges Wissen repräsentiert werden kann.

Zunächst sei angenommen, dass es Operatoren gibt, die SRL-Operatoren als Argumente haben. Dann lassen sich die Regeln für die Prädikate, die Transitivität und Konversität repräsentieren, folgendermaßen darstellen, wobei p und q Operatoren der SRL seien.

$$\begin{array}{ll} \text{Wenn Transitiv}(p), \text{ dann gilt:} & p(x,y) \wedge p(y,z) \Rightarrow p(x,z) \\ \text{Wenn Konvers}(p,q), \text{ dann gilt:} & p(x,y) \Rightarrow q(y,x) \end{array}$$

Hierbei ist zu beachten, dass es sich immer noch um die normale logische Implikation, also normale Regeln handelt. Nur die Operatoren $\text{Transitiv}(x)$ und $\text{Konvers}(x)$ sind höhere Konzepte, da sie Operatoren als Argumente haben. Der nächste (Generalisierungs-) Schritt ist, Regeln über die Meta-Operatoren zu erstellen. Im Fall des hier eingeführten Beispiels kann ein solcher Meta-Operator folgendermaßen aussehen, wobei das M über dem Implikationszeichen zum Ausdruck bringen soll, dass es sich um einen Metajunktoren handelt.

$$\text{Transitiv}(p) \wedge \text{Konvers}(p,q) \stackrel{M}{\Rightarrow} \text{Transitiv}(q)$$

Die Wissensbasis eines Agenten besteht aus einer Menge von geschlossenen Formeln (Menge der Fakten F) und einer Menge von offenen Formeln (Menge der Regeln R). Nimmt man eine Wissensbasis als gegeben an, ist das zu einem Operator oder Konzept gehörende Wissen nun durch folgende interne Struktur gegeben:

¹⁹ Regeln in SRL können auch als Default-Regeln dargestellt sein. Das bedeutet, dass Schlüsse und die dadurch zur WB hinzugefügte Fakten, zu einem späteren Zeitpunkt wieder rückgängig gemacht werden können. Trotzdem sollte es nach [Hab86] einem WBS möglich sein, zum jeweiligen Zeitpunkt mit seiner WB so zu arbeiten, als ob diese nur sicheres Wissen enthält. Für eine Einführung in Defaultlogik vgl. z. B. [Rei91] und [RuN03].

- Operatorname: op
- Wohlgeformtheitseigenschaften und –bedingungen:
 - sorte (op)
 - arg-sorten (op)
- Konzeptuelle Beziehungen und Eigenschaften:
 - Regeln, in denen op als Prämisse oder Konklusion auftritt
 - Metafakten, in denen op als Argument auftritt
- Weltwissen über ‚ op ‘:
 - Fakten, d.h. $p \in F$, in denen ‚ op ‘ auftritt.

Dies gilt natürlich sowohl für SRL-Operatoren als auch für Meta-SRL-Operatoren. Die Einführung von Regeln sowie von Metaregeln ist Abstraktion. Die Notwendigkeit, solche Regeln in einer WB zu repräsentieren, liegt auf der Hand, wenn man eine WB als kognitives Modell eines Menschen betrachtet.

3.1.2 Referentielle Netze

Bei der folgenden überblickartigen Beschreibung referentieller Netze werde ich mich an die in [Hab88] gegebene Spezifikation halten, aber nur die m. E. zum Verständnis der folgenden Abschnitte notwendigen Teile der Theorie wiedergeben. Die Grundstruktur von referentiellen Netzen enthält Referenzobjekte, Designationen und Attribute von Referenzobjekten.

Definition 3.1.3: *Die Menge REFO ist die Menge der Referenzobjekte, die Menge DESIGN die der Designationen und die Menge R-ATT die der Attribute von Referenzobjekten.*

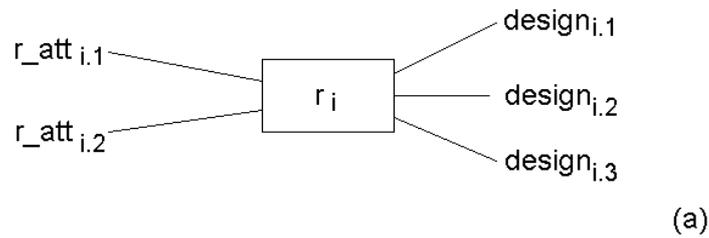
Das Definitionsschema 3.1.2, das Terme und Formeln der SRL induktiv definiert, muss nun so erweitert werden, dass eine zusätzliche Menge von Termen existiert. Diese zusätzliche Menge ist die Menge der Referenzobjekte, was durch folgende Definition zum Ausdruck gebracht wird.

Definition 3.1.4: *SRL-Terme fallen in eine der folgenden Kategorien:*

- Referenzobjekte (REFO)
- Variablen (VAR)
- Name (NAM, Operatoren op mit $typ(op) = \langle t, 0, 0, 0 \rangle$)
- Deskriptionen (DESKR, Operatoren op gebildet mit $typ(op) = \langle t, n, m, p \rangle$)

Die Menge DESIGN der Designationen ist eine Teilmenge der Vereinigungsmenge der Namen und der geschlossenen Deskriptionen.

$$DESIGN \subset NAM \cup DESKR_{cl}.$$



$$(R-ATT \quad X \quad REFO) \quad X \quad DESIGN$$

(b)

Abbildung 3.1: (a) Graphische Veranschaulichung eines Referenzobjekts. (b) Struktur eines referentiellen Netzes.

Referenzobjekte fungieren in referentiellen Netzen als Stellvertreter von Objekten der Welt²⁰ bzw. des Denkens. Die Eigenschaften oder Merkmale nach [Rei91] (vgl. oben) der Referenzobjekte werden durch die Designationen und die Attribute repräsentiert. Es wird noch die Möglichkeit benötigt, essentielle Eigenschaften²¹ wie zum Beispiel Sorten im Netz zu repräsentieren. Diese essentiellen Eigenschaften werden mittels der Attribute repräsentiert. Gegeben eine nicht leere, höherdimensionale Menge von Attributen von Referenzobjekten R-ATT und eine Abbildung $r\text{-att}: REFO \rightarrow R\text{-ATT}$ ²². Dann gilt für ein referentielles Netz folgendes:

$$RNET \subset (R-ATT \times REFO) \times DESIGN$$

Die Attribute werden von Habel durch Vektoren dargestellt, was durch den Begriff höherdimensional ausgedrückt wird. Die Klammer um R-ATT und REFO soll andeuten, dass die Referenzattribute funktional von den Referenzobjekten abhängen.

²⁰ Objekte oder Konzepte beziehen sich natürlich auf die projizierte Welt (vgl. Abschnitt 2.1). Ich werde im Folgenden von der Welt sprechen und dabei die projizierte Welt meinen.

²¹ Für eine Erklärungen, was essentiell bedeuten soll oder bedeuten kann vgl. [Hab86] S.163ff.

²² Die Funktion r-att in den referentiellen Netzen ist sinngemäß vergleichbar mit der Funktion sor-z in der SRL.

3.2 Taxonomien

In diesem Abschnitt werden nun Taxonomien formell eingeführt. Taxonomien sind Vererbungshierarchien, wie sie im Zusammenhang mit der Bestimmung der Ähnlichkeit von Konzepten bereits in Abschnitt 1.2 und im vorangegangenen Abschnitt für Sortenhierarchien vorgestellt wurden. Generell soll in Taxonomien Konzeptwissen repräsentiert werden. Vererbt werden die in einem Konzept enthaltenen Informationen.

Konzepte können spezifischer oder allgemeiner als andere Konzepte sein oder einander ausschließen. So sind zum Beispiel die Konzepte Mensch und Vogel spezifischer als das Konzept Lebewesen, aber die Konzepte Mensch und Vogel schließen einander aus. Das Konzept Lebewesen subsumiert die Konzepte Mensch und Vogel, was ausdrückt, dass sowohl Menschen als auch Vögel Lebewesen sind. Der gegenseitige Ausschluss (Exklusion) der Konzepte Mensch und Vogel bringt zum Ausdruck, dass ein Mensch niemals ein Vogel sowie umgekehrt ein Vogel niemals ein Mensch ist. Die Exklusion vererbt sich über die Subsumption. Das bedeutet zum Beispiel, dass alle vom Konzept Mensch subsumierten Konzepte wie etwa das Konzept Kind ebenfalls niemals zum Konzept Vogel gehören.

Eine Weiterführung dieses Gedankens bilden die so genannten Beschreibungslogiken, in denen explizit eine Syntax und Semantik im Sinne einer Logik definiert werden und mit denen sehr komplexe Konzeptbeschreibungen vorgenommen werden können. Vorläufer von Beschreibungslogiken ist zum Beispiel das von Schmolze und Lipkis in [ScL83] vorgestellte KL-ONE-System. Beispiele für Beschreibungslogiken sind DLP und Fact. Für eine Einführung in das Feld der Beschreibungslogiken sei auf [BCM03] verwiesen. Die Verwendung einer Beschreibungslogik ist für die in Kapitel 4 vorgestellten Bestimmungen von Ähnlichkeitsmaßen nicht notwendig.

In der vorliegenden Arbeit werden Taxonomien verwendet, um die Ähnlichkeit von Konzepten zu bestimmen. Wie in Abschnitt 1.2.1 gefordert wurde, soll sich die Ähnlichkeit von Repräsentationen nur auf Informationen beziehen, die in den verglichenen Repräsentationen enthalten ist. Es muss also bestimmt werden können, welche Informationen zwei Konzepte gemeinsam haben. Die Subsumptionsrelation bringt zum Ausdruck, dass ein Konzept c , das von einem Konzept c' subsumiert wird, mindestens alle Eigenschaften von c' hat. Für die in c enthaltene Information bedeutet dies, dass sie eine Teilmenge der Information ist, die in c' enthalten ist.²³

Eine Taxonomie besteht nun aus einer Menge von Konzeptsymbolen, sowie der Subsumptionsrelation und der Exklusionsrelation. Die Subsumptionsrelation ist reflexiv, antisymmetrisch und transitiv. Die Exklusionsrelation ist

²³ In Taxonomien wird häufig Default-Wissen repräsentiert, also Wissen, das normalerweise zutrifft. Defaultwissen soll in dieser Arbeit nicht behandelt werden. Eine Einführung in Default-Logik wird zum Beispiel in [RuN03] gegeben.

irreflexiv und wird über die Subsumption vererbt. Weiterhin ist in der Menge der Symbole mindestens das Symbol T (gesprochen „Top“) enthalten. Das Symbol T repräsentiert das allgemeine Konzept, also das Konzept, das alle anderen Konzepte subsumiert. Besteht zwischen zwei Konzepten c_1 und c_2 die Subsumptionsrelation, so wird im Folgenden von einer Vater-Kind-Beziehung zwischen c_1 und c_2 gesprochen. Besteht zwischen je einem der beiden Konzepte und einem dritten Konzept c_3 eine Vater-Kind-Beziehung, so wird im Folgenden von c_1 und c_2 als Brüder gesprochen.

Definition 3.2.1: Eine Taxonomie $T = (S, \leq, \leftrightarrow)$ besteht aus einer endlichen Menge S von Symbolen der Konzepte, die mindestens das Symbol T enthält, einer reflexiven, antisymmetrischen und transitiven Relation $\leq \subset S \times S$, genannt Subsumptionsrelation, und einer irreflexiven Relation $\leftrightarrow \subset S \times S$, genannt Exklusionsrelation.

Für $(s_1, s_2) \in \leq$ wird auch $s_1 \leq s_2$ geschrieben und für $(s_1, s_2) \in \leftrightarrow$ wird auch $s_1 \leftrightarrow s_2$ geschrieben. Es liegt eine direkte Subsumption zwischen s_1 und s_2 vor, wenn $s_1 \leq s_2$ und es gilt, dass es kein Symbol $s \in S$ gibt, für das gilt $s_1 \leq s \leq s_2$. Außerdem gilt folgendes:

- a) $\forall s \in S: s \leq T$
- b) $\forall s_1, s_2 \in S: s_1 \leftrightarrow s_2 \Rightarrow (\forall s \in S: s \leq s_1 \Rightarrow s \leftrightarrow s_2) \wedge (\forall s \in S: s \leq s_2 \Rightarrow s \leftrightarrow s_1)$

Die Antisymmetrie der Subsumptionsrelation ist wichtig, da es keine Zyklen in der Subsumption geben darf. So ist zum Beispiel für die Konzepte Mensch, Lebewesen und Kind der folgende Ausdruck nicht sinnvoll:

$$\text{Mensch} \leq \text{Lebewesen} \wedge \text{Kind} \leq \text{Mensch} \wedge \text{Lebewesen} \leq \text{Kind}$$

Solche Zyklen sind durch die Antisymmetrie der Subsumptionsrelation ausgeschlossen.

Um ein Maß der Ähnlichkeiten zwischen Konzepten zu bestimmen (vgl. Kapitel 4), müssen noch einige Taxonomien betreffende Begriffe bestimmt werden. Von entscheidender Bedeutung sind die Konzepte, die beide verglichene Konzepte subsumieren. Sie repräsentieren Konzepte der Taxonomie, die nur Informationen enthalten, die auch in beiden verglichenen Konzepten enthalten sind. Offensichtlich subsumiert T zwei beliebige verglichene Konzepte. Allerdings sagt T nichts über die Information aus, die in beiden Konzepten enthalten ist, da die in T enthaltene Menge von Informationen leer ist. Es soll deshalb ein Konzept der Taxonomie bestimmt werden, das die verglichenen Konzepte subsumiert, aber wiederum kein Konzept subsumiert, welches ebenfalls beide Konzepte subsumiert. Das bedeutet, dass das so bestimmte Konzept ein Maximum der Informationen enthält, die in beiden Konzepten enthalten ist. In Anlehnung an die

Bezeichnung der Vater-Kind-Beziehung soll dieses Konzept als spezifischster Vater bezeichnet werden.

Definition 3.2.2: Seien c, c_1 und c_2 Elemente einer Taxonomie $T = (S, \leq, \leftrightarrow)$. Dann ist c ein spezifischster Vater von c_1 und c_2 , wenn c ein Konzept aus S ist, dass sowohl c_1 als auch c_2 subsumiert, und für alle Konzepte c' aus S , die sowohl c_1 als auch c_2 subsumieren, gilt, dass c von c' subsumiert wird.

$$c = \text{spezVater}_T(c_1, c_2) \stackrel{\text{def}}{\Leftrightarrow} c_1 \leq c \wedge c_2 \leq c \wedge \forall c' \in S: [c_1 \leq c' \wedge c_2 \leq c' \Rightarrow c \leq c']$$

Dass es für zwei Konzepte einer Taxonomie mindestens einen spezifischsten Vater geben muss, ergibt sich direkt aus Definition 3.2.1 a).

Für die Bestimmung der Ähnlichkeit von zwei Konzepten muss bekannt sein, wie weit die Konzepte in der Taxonomie voneinander entfernt sind. Eine Möglichkeit, die Entfernung der Konzepte voneinander zu bestimmen, ist, die Pfadlänge in einem Graphen zu bestimmen, dessen Kanten durch die Subsumptionsrelation bestimmt werden. Allerdings ist die Subsumption transitiv und reflexiv. Diese Eigenschaft führt dazu, dass die Pfadlänge zwischen zwei identischen Konzepten einer Taxonomie gleich eins ist und die Pfadlänge zwischen einem Konzept c und einem von c subsumierten Konzept ebenfalls gleich eins ist. Dies ist offensichtlich nicht wünschenswert. Vielmehr wird für die Bestimmung der Pfadlänge der Graph benötigt, dessen Kanten die direkte Subsumption repräsentieren und in dem reflexive Kanten verboten sind. Um den Subsumptionsgraphen zu definieren, soll zunächst definiert werden, was im Folgenden unter einem Graphen verstanden wird.

Definition 3.2.3: Ein gerichteter Graph $G = (V, E)$ ist ein Paar aus einer endlichen Menge $V \neq \emptyset$ und einer Menge E von geordneten Paaren (a, b) mit $a \neq b$ aus V . Die Elemente von V heißen **Knoten**, die Elemente von E heißen **Kanten**.

Für ein Element (a, b) aus E heißt a der **Anfangsknoten** und b der **Endknoten** der Kante.

Im Folgenden soll unter einem Graphen immer ein gerichteter Graph nach Definition 3.2.3 verstanden werden. Nun wird noch ein Pfad eines Graphen definiert.

Definition 3.2.4: Für einen Graphen $G = (V, E)$ heißt eine Folge von Kanten (e_1, \dots, e_n) gerichteter **Pfad**, wenn die e_i für $i = 1, \dots, n$ paarweise verschieden sind und es Knoten v_0, \dots, v_n gibt mit $e_i = (v_{i-1}, v_i)$ für $i = 0, \dots, n$ und die v_j für $j = 0, \dots, n$ paarweise verschieden sind.

Für $e_1 = (v_0, v_1)$ und $e_n = (v_{n-1}, v_n)$ heißt v_0 **Anfangsknoten** des gerichteten Pfades und v_n **Endknoten** des gerichteten Pfades.

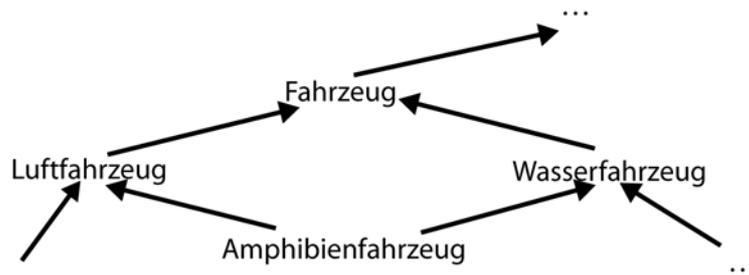


Abbildung 3.1: Beispiel eines Ausschnitts aus einem Subsumptionsgraphen.

Da in der vorliegenden Arbeit nur gerichtete Pfade verwendet werden, soll im Folgenden unter einem Pfad immer ein gerichteter Pfad verstanden werden. Auf eine formelle Einführung in die Theorie von Graphen wird hier verzichtet und auf [Jun94] verwiesen. Nun kann der Subsumptionsgraph definiert werden.

Definition 3.2.5: Für eine Taxonomie $T = (S, \leq, \leftrightarrow)$ ist der Subsumptionsgraph $suG(T) = (V, E)$ ein gerichteter Graph, der definiert ist durch die Menge der Knoten

$$V = S$$

und die Menge der Kanten

$$E = \{ (s_1, s_2) \in \leq_A \mid s_1 \neq s_2 \wedge \neg \exists s \in S_A: s_1 \leq_A s \leq_A s_2 \}.$$

Der gesuchte Pfad im Subsumptionsgraphen soll für die Bestimmung der Entfernung zwischen zwei Konzepten und die Entfernung für die Bestimmung der Ähnlichkeit von den Konzepten verwendet werden. Bei der Bestimmung der Ähnlichkeit von Konzepten darf die Forderung, dass sich die Ähnlichkeit von Repräsentationen auf Informationen beziehen soll, die in beiden Repräsentationen enthalten ist (vgl. Abschnitt 1.2.1), nicht verletzt werden. Um dies zu garantieren, bietet es sich an, zwei Pfade für die Bestimmung der Entfernung zwischen zwei Konzepten zu verwenden, je ein Pfad zwischen einem der Konzepte und dem spezifischsten Vater der Konzepte. Deshalb sei zunächst der Pfad zwischen zwei Konzepten einer Taxonomie definiert.

Definition 3.2.4: Gegeben sei eine Taxonomie T und ihr Subsumptionsgraph $suG(T) = (V, E)$. Ein Pfad zwischen zwei Konzepten c_1 und c_n aus T ist eine Folge von Elementen aus E , für die Folgendes gilt:

$$\text{pfad}_T(c_1, c_n) \stackrel{\text{def}}{=} \{ (c_1, c_2), \dots, (c_{n-1}, c_n) \mid (c_x, c_{x+1}) \in E, x = 1, \dots, n-1 \}$$

Dass die Bestimmung der Entfernung zwischen zwei verglichenen Konzepten einen spezifischsten Vater mit einbeziehen sollte, ist darin begründet, dass ein

spezifischster Vater nur Informationen enthält, die in beiden verglichenen Konzepten enthalten sind. Anstelle eines spezifischsten Vaters könnte auch ein Konzept verwendet werden, das den spezifischsten Vater subsumiert. Dann wäre zwar die Forderung aus Abschnitt 1.2.1 erfüllt, die so bestimmte Ähnlichkeit wäre aber zu klein. Die Verwendung der gerichteten Pfade dient ebenfalls der Erfüllung der Forderung. Durch die Verwendung von gerichteten Pfaden zum spezifischsten Vater ist garantiert, dass der Pfad nur aus Kanten besteht, deren Endpunkte eine Teilmenge der Information des Anfangsknotens enthalten. Also enthält jedes auf dem Pfad gelegene Konzept eine Teilmenge der Information des Anfangsknotens des Pfades. Verwendet man also solche Pfade, ist garantiert, dass für die Bestimmung der Ähnlichkeit nur Werte verwendet werden, die durch Teilmengenbildung der Informationen der verglichenen Konzepte bestimmt wurden. Somit ist die Forderung aus Abschnitt 1.2.1 erfüllt.

Folgendes Beispiel soll dies anhand des in Abbildung 3.1 dargestellten Subsumptionsgraphen verdeutlichen. Angenommen, es soll die Ähnlichkeit zwischen den Konzepten Landfahrzeug und Wasserfahrzeug festgestellt werden und es sind beliebige Pfade für die Bestimmung der Entfernung zwischen Konzepten erlaubt. Dann kann auch der Pfad gewählt werden, der das Amphibienfahrzeug enthält. Es ist dann aber nicht feststellbar, welche der Informationen aus dem Konzept Amphibienfahrzeug sowohl im Konzept Wasserfahrzeug als auch im Konzept Landfahrzeug enthalten sind. Angenommen, das Konzept Amphibienfahrzeug enthält die Information, dass es sich um ein Fahrzeug handelt, das sich an Land bewegen kann und das sich im Wasser bewegen kann. Zwar ist ein Teil dieser Informationen auch in den beiden verglichenen Konzepten enthalten, welche dieser Informationen sowohl im Land- als auch im Wasserfahrzeug enthalten ist, kann aber nicht festgestellt werden. Verwendet man wie eben beschrieben die Pfade zwischen den verglichenen Konzepten und dem spezifischsten Vater (das Fahrzeug), ist klar, welche Information sowohl im Landfahrzeug als auch im Wasserfahrzeug enthalten ist, nämlich dass es sich bei beiden Konzepten um Fahrzeuge handelt. Für die Bestimmung der Entfernung zwischen zwei Konzepten kann nun die Länge der Pfade zwischen den Konzepten und deren spezifischsten Vater addiert werden. Zwischen den verglichenen Konzepten und ihrem spezifischsten Vater kann es mehrere Pfade geben. Es sollen die kürzesten Pfade für die Bestimmung der Entfernung verwendet werden. Da der spezifischste Vater nicht eindeutig ist, hat die folgende Funktion zur Bestimmung der Entfernung zweier Konzepte einer Taxonomie als drittes Argument einen spezifischsten Vater.

Definition 3.2.5: Für die Konzepte c_1 , c_2 und c_{SV} einer Taxonomie T , mit $c_{SV} = \text{spezVater}_T(c_1, c_2)$ ist die Entfernung zwischen c_1 und c_2 die Summe aus der Anzahl der Elemente des kürzesten Pfads zwischen c_1 und c_{SV} und der Anzahl der Elemente des kürzesten Pfads zwischen c_2 und c_{SV} .

$$\text{entfernung}_T(c_1, c_2, c_{SV}) = \begin{cases} \min \{ (|\text{pfad}_T(c_1, c_{SV})| + |\text{pfad}_T(c_2, c_{SV})|) \}, & \text{wenn } c_{SV} = \text{spezVater}(c_1, c_2) \\ \text{undefiniert} & , \text{sonst} \end{cases}$$

Für die in Kapitel 4 vorgestellten Ähnlichkeitsmaße muss noch die Tiefe eines Konzepts im Subsumptionsgraphen bestimmt werden können. Sie ist die kleinste Pfadlänge zwischen dem Konzept und T .

Definition 3.2.6: Die Tiefe eines Konzepts c einer Taxonomie T ist die Anzahl der Elemente des kürzesten Pfades zwischen c und T .

$$\text{tiefe}_T(c) = \min \{ |\text{pfad}_T(c, T)| \}$$

Für die Bestimmung der Ähnlichkeit von Konzepten muss noch die maximale Entfernung von Konzepten einer Taxonomie bestimmt werden können.

Definition 3.2.7: Für eine Taxonomie $T = (S, \leq, \leftrightarrow)$ ist die maximale Entfernung bestimmt durch

$$\text{maxPL}(T) = \max \{ \text{entfernung}_T(c_1, c_2, T) \mid (c_1, c_2) \in S \times S \}.$$

Die Entfernung wird wie gesagt für die Bestimmung der Ähnlichkeit zweier Konzepte benötigt. Dass die Entfernung zu diesem Zweck verwendet werden kann, soll nun noch einmal deutlich gemacht werden. Generell spricht eine kleine Entfernung zwischen zwei Konzepten für eine hohe Ähnlichkeit und umgekehrt eine große Entfernung für eine kleine Ähnlichkeit. Die Abhängigkeit der Ähnlichkeit zweier Konzepte von ihrer Entfernung in der Taxonomie erscheint einleuchtend, wenn man sich die Struktur der Taxonomien anschaut. Generell impliziert eine große Entfernung eine große Anzahl an Spezialisierungen und/oder Generalisierungen. Jede Generalisierung bedeutet zunächst einen Informationsverlust, eine Spezialisierung einen Informationsgewinn. Das bedeutet, dass auf dem Pfad zum spezifischsten Vater der beiden Konzepte zunächst bestimmte Informationen verloren gehen und dann auf dem Pfad vom spezifischsten Vater zum anderen Konzept *andere* Informationen gewonnen werden. Wenn der spezifischste Vater nicht T ist, haben die beiden Konzepte einen gemeinsamen Anteil an Informationen, der im spezifischsten Vater enthalten ist. Der Anteil an verschiedener Information wird jedoch mit zunehmender Entfernung voneinander größer, da eine zunehmende Pfadlänge bedeutet, dass entweder ein größerer Teil an

Information während der Generalisierung verloren geht oder ein größerer Teil an verschiedener Information während der Spezialisierung hinzu kommt oder beides. Also wird die Ähnlichkeit zweier Konzepte mit zunehmender Entfernung geringer, da es immer mehr Information gibt, in der sich die Konzepte unterscheiden.

3.3 CRIL-Graphen

In diesem Abschnitt wird ein formelles Modell von CRIL-Graphen vorgestellt. CRIL-Graphen haben die gleiche Struktur wie die in Abschnitt 3.1 vorgestellten referentiellen Netze. So ist es möglich, ein referentielles Netz auf einen CRIL-Graphen abzubilden und umgekehrt. Dies ist notwendig, da für die syntaktische und semantische Analyse der Wegbeschreibung sowie die Repräsentation der Satzbedeutung referentielle Netze verwendet werden. Außerdem sollen in referentielle Netzen repräsentierte Regeln auch auf CRIL-Graphen anwendbar sein. Es muss deshalb möglich sein, das durch referentielle Netze repräsentierte Wissen durch einen CRIL-Graphen zu repräsentieren. Der in Abschnitt 2.2 vorgestellte Instruktionsgraph des Instruktionsmodells ist ein CRIL-Graph.

Ein CRIL-Graph besteht aus einer Menge von CRIL-Knoten und (gerichteten) Kanten zwischen den Knoten. Die Kanten des CRIL-Graphen repräsentieren mehrstellige Operatoren, also Relationen eines referentiellen Netzes. Da in referentiellen Netzen auch Operatoren erlaubt sind, die mehr als zweistellig sind, muss der CRIL-Graph offensichtlich ein gerichteter Hypergraph sein. Um zu repräsentieren, welchen Operator eine Kante repräsentiert, gibt es eine Kantenfunktion, die Kanten auf Operatorensymbole abbildet. Jeder CRIL-Knoten repräsentiert ein Referenzobjekt. Sorten, einstellige Prädikate und Namen des Referenzobjekts werden im CRIL-Knoten durch Mengen entsprechender Symbole repräsentiert. So lässt sich ein referentielles Netz auf einen CRIL-Graphen sowie umgekehrt ein CRIL-Graph auf ein referentielles Netz abbilden.

Für den Vergleich von CRIL-Graphen werden noch Taxonomien benötigt. Zum einen wird eine Taxonomie benötigt, die den Sortenhierarchien der referentiellen Netze entspricht. In ihr sollen auch die Konzepte enthalten sein, die in referentiellen Netzen durch einstellige Operatoren repräsentiert werden. Diese Taxonomie heißt die Taxonomie der Attribute. Zum anderen wird eine Taxonomie der Konzepte benötigt, die in referentiellen Netzen durch mehrstellige Operatoren, also Relationen repräsentiert sind. Diese explizite Trennung bietet sich aus zwei Gründen an. Zum einen ist die Menge der Relationensymbole offensichtlich der Wertebereich der Kantenfunktion. Zum anderen werden die Taxonomien während des Vergleichs von CRIL-Graphen in unterschiedlichen Kontexten verwendet, wie in Kapitel 4 noch deutlich werden wird.

Definition 3.3.1: Gegeben sei die Taxonomie $T_A = (S_A, \leq_A, \leftrightarrow_A)$ der Attribute und die Taxonomie $T_R = (S_R, \leq_R, \leftrightarrow_R)$ der Relationen.

Ein CRIL-Knoten $ck = (SOR, EIG, NAM)$ ist ein Tripel, bestehend aus einer Menge SOR von Sorten, für die gilt $SOR \subseteq S_A$, einer Menge EIG von Eigenschaften, für die gilt $EIG \subseteq S_A$ sowie einer Menge NAM , deren Elemente Namen, also Konstanten repräsentieren. Es werden ferner die folgenden Funktionen über einen CRIL-Knoten definiert.

$$\begin{aligned} \mathcal{S}(ck) &= SOR \\ \mathcal{E}(ck) &= EIG \\ \mathcal{N}(ck) &= NAM \end{aligned}$$

Ein CRIL-Graph $G = (V, E, T_A, T_R, op)$ besteht aus der Menge V von CRIL-Knoten, der Menge E der Kanten $E \subseteq (K^2 \cup K^3 \cup \dots)$, den Taxonomien T_A und T_R und der Operatorfunktion über die Kanten $op: E \rightarrow S_R$. T_E und T_R heißen die Taxonomie der Attribute und die Taxonomie der Relationen von G . Es werden ferner folgende Definitionen über CRIL-Graphen definiert.

$$\begin{aligned} \mathcal{V}(G) &= V \\ \mathcal{E}(G) &= E \end{aligned}$$

Die beiden Taxonomien aus Definition 3.3.1 gehören zum Weltwissen des Agenten. Sie werden sowohl in der Instruktionsphase für die syntaktische und semantische Analyse als auch in der Navigationsphase des GA verwendet. Die referentiellen Netze, die die Satzbedeutung repräsentieren, verwenden also die gleichen Taxonomien und somit auch die gleichen Konzepte. So ist es ohne weiteres möglich, die referentiellen Netze auf CRIL-Graphen abzubilden. Für die Bestimmung der Ähnlichkeit von CRIL-Knoten muss es möglich sein, zu unterscheiden, ob ein CRIL-Knoten Wissen der Wegbeschreibung oder der Perzeption repräsentiert. Deshalb soll folgende Definition gegeben werden.

Definition 3.3.2: Ein I-Graph ist ein CRIL-Graph, der Wissen einer Wegbeschreibung repräsentiert. Seine Knoten heißen I-Knoten, wenn sie Wissen der Instruktion repräsentieren und P-Knoten, wenn sie Wissen der Perzeption repräsentieren. Alle im I-Graphen enthaltenen P-Knoten müssen mit einer Koreferenzkante mit einem I-Knoten verbunden sein.

Ein P-Graph ist ein CRIL-Graph, der Wissen repräsentiert, das ein Agent durch Perzeption gewonnen hat. Seine Knoten heißen P-Knoten.

Die Taxonomie der Attribute und die Taxonomie der Relationen des I-Graphen müssen auch die entsprechenden Taxonomien des P-Graphen sein.

Ein vom Agenten wahrgenommenes Objekt der Umgebung wird also durch einen P-Knoten repräsentiert. Der Knoten repräsentiert aber nur bestimmte

Attribute des Objekts sowie eventuell Relationen des Objekts zu anderen Objekten. Diese Repräsentation reicht aber nicht unbedingt aus, um das Objekt zu einem späteren Zeitpunkt zu identifizieren. So kann es vorkommen, dass der Agent ein Haus wahrnimmt und das Haus durch einen Knoten repräsentiert, der nur das Attribut Haus hat. Wenn es in der Nähe des Hauses noch andere Häuser gibt, reicht die Eigenschaft, ein Haus zu sein, offensichtlich nicht aus, um das Haus wieder zu erkennen. Ein Mensch in der gleichen Situation wäre aber in der Lage, das Haus wieder zu erkennen. Nun wäre es gerechtfertigt anzunehmen, dass der Mensch sich genügend Attribute gemerkt hat, um das Haus eindeutig zu identifizieren, also zum Beispiel welche Farbe oder wie viele Fenster oder welche anderen hervorstechende Attribute das Haus hat. Die Tatsache, dass es einem Menschen aber erst nach längerem Überlegen möglich ist, bestimmte Attribute zu nennen, durch die das Haus eindeutig identifizierbar wäre, deutet aber darauf hin, dass der Mensch sich eben nicht nur Attribute des Hauses merkt, sondern eher ein Bild des Hauses. Dieses Bild entspricht der in Abschnitt 1.1.1 erläuterten Sicht, die für die Identifikation von Plätzen in kognitiven Karten verwendet wird.

In einer virtuellen Umgebung wie der des Geometrischen Agenten sind außerdem nicht unbedingt genug Attribute der Objekte repräsentiert, um nur anhand der Attribute eines Objekts eine eindeutige Identifizierung des Objekts zu garantieren. Damit der Agent in der Lage ist, Objekte der Umgebung wieder zu erkennen, muss es also einen zusätzlichen Mechanismus geben. Wie dieser Mechanismus auszusehen hat, sei hier nicht bestimmt. In einer simulierten Umgebung wie der des Geometrischen Agenten ist z. B. eine direkte Referenz auf das Objekt denkbar. Für Agenten in realen Umgebungen wären dagegen Momentaufnahmen denkbar, anhand derer das Objekt später identifiziert werden kann. Diese Momentaufnahmen entsprechen dann wiederum Sichten. Wichtig ist, dass das Objekt zu einem späteren Zeitpunkt wieder erkannt werden kann.

Kapitel 4

Ähnlichkeitsmaße

Die im Folgenden vorgestellte Bestimmung von Ähnlichkeiten soll es einem Agenten während der Navigation ermöglichen zu bestimmen, ob ein I-Knoten und ein P-Knoten koreferent sind, also das gleiche Objekt repräsentieren. Dabei sollen Relationen zwischen den repräsentierten Objekten ebenfalls berücksichtigt werden. Folgendes Beispiel soll dies anhand Abbildung 4.1 verdeutlichen. Angenommen im I-Graphen sind zwei Knoten I und I' enthalten sowie eine Relation²⁴ zwischen I und I' . Im P-Graphen sind drei Knoten P_1 , P_2 und P' enthalten. P_1 und P_2 sollen beide große Ähnlichkeit mit I haben und P' soll große Ähnlichkeit mit I' haben. P_1 und P' stehen in der gleichen Relation zueinander wie I und I' . In diesem Beispiel sollte der Agent bei der Identifikation der Koreferenz in der Lage sein zu erkennen, dass P_1 und I sowie P' und I' koreferent sind, da die Knoten nicht nur große Ähnlichkeit haben, sondern auch in der gleichen Relation zueinander stehen. Angenommen, es sei möglich, ein Maß der Ähnlichkeit von Knoten und

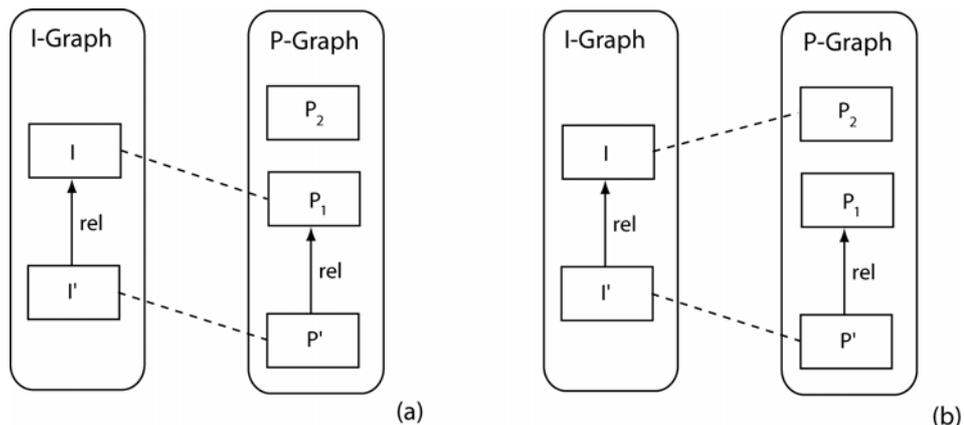


Abbildung 4.1: Koreferenzen (gestrichelte Kante) zwischen einem I-Graphen und einem P-Graphen. (a) Korrekte Koreferenz unter Berücksichtigung der Relation. (b) Inkorrekte Koreferenz.

²⁴ Ich werde im Folgenden anstelle von einer Kante, für die die Kantenfunktion auf das Konzept einer Relation abbildet, abkürzend von einer Relation sprechen.

Kanten zweier Graphen zu bestimmen. Die Ähnlichkeit soll durch Verknüpfungen der Knoten und Kanten der beiden Graphen repräsentiert sein. Dann ist die Summe der Ähnlichkeiten in Abbildung 4.1(a) offensichtlich größer als die der Verknüpfung in Abbildung 4.1(b). Es muss also eine Verknüpfung der Elemente der Graphen gefunden werden, die die Ähnlichkeit der Graphen bezüglich der Verknüpfung maximiert. Wie diese Verknüpfung gefunden werden kann, wird in Abschnitt 4.5 gezeigt. Um diese Verknüpfung zu finden, muss es wie gesagt möglich sein, die Ähnlichkeit von Konzepten, zum Beispiel Sorten oder Relationen, von ganzen Knoten und von einer gegebenen Verknüpfung zweier Graphen zu bestimmen. Wie diese Ähnlichkeiten bestimmt werden können, wird in den Abschnitten 4.2 bis 4.4 gezeigt. Im folgenden Abschnitt 4.1 soll zunächst allgemein auf die Ähnlichkeit zweier Repräsentationen eingegangen werden.

4.1 Ähnlichkeit von Repräsentationen

Zunächst soll betrachtet werden, wie im Allgemeinen ein Maß für die Ähnlichkeit zweier Repräsentationen bestimmt werden kann. Ein Maß der Ähnlichkeit von Repräsentationen soll von den Informationen abhängen, die in beiden Repräsentationen enthalten sind. Poole und Campbell in [PoC95] bestimmen dafür eine Funktion `informationsgehalt`, die ein Maß für die Information berechnet, welche in einer Repräsentation enthalten ist.²⁵ Gegeben seien zwei beliebige Repräsentationen A und B sowie eine Funktion `gemInfo(A,B)`, die eine Repräsentation berechnet, welche nur Informationen enthält, die sowohl in A als auch in B enthalten sind. Dann kann ein Maß für die Ähnlichkeit²⁶ zwischen A und B durch folgende, aus [PoC95] stammende Gleichung bestimmt werden.

Gleichung 4.1.1: *Ein Ähnlichkeitsmaß für zwei beliebige Repräsentationen A und B ist gegeben durch*

$$\text{ähnlichkeit}(A,B) = \frac{\text{informationsgehalt}(\text{gemInfo}(A,B))}{\max(\text{informationsgehalt}(A), \text{informationsgehalt}(B))}$$

Alternativ zu Gleichung 4.1.1 wird in der Literatur (vgl. z.B. [MGL00a]) häufig der Dice-Koeffizient benutzt, was zu folgender Gleichung führt.

²⁵ Dabei soll an dieser Stelle noch nicht spezifiziert sein, wie dieses Maß bestimmt wird.

²⁶ Im Folgenden wird für das Maß für die Ähnlichkeit nur noch der Ausdruck Ähnlichkeit verwendet.

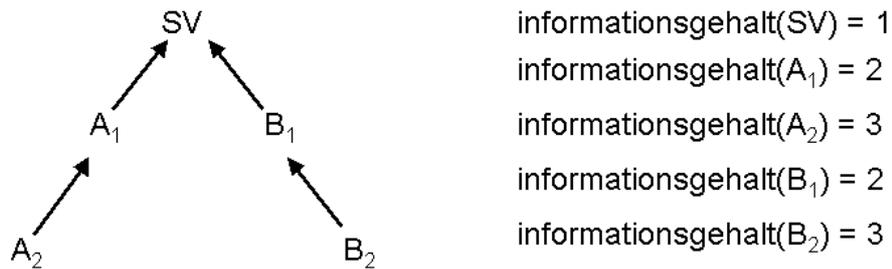


Abbildung 4.2: Beispieltaxonomie von Repräsentationen mit ihren Informationsgehalten.

Gleichung 4.1.2: Ein Ähnlichkeitsmaß für zwei beliebige Repräsentationen *A* und *B* ist gegeben durch

$$\text{ähnlichkeit}(A,B) = \frac{2 \text{ informationsgehalt}(\text{gemInfo}(A,B))}{\text{informationsgehalt}(A) + \text{informationsgehalt}(B)}$$

Die durch Gleichungen 4.1.1 und 4.1.2 bestimmten Ähnlichkeiten haben einen Wertebereich zwischen 0 und 1. Die so bestimmte Ähnlichkeit ist 0 für Repräsentationen, die keinerlei gemeinsame Informationen haben, und 1, für identische Repräsentationen. Dass die Gleichungen 4.1.1 und 4.1.2 für Repräsentationen, die keine gemeinsame Information enthalten 0 ergeben, ergibt sich direkt aus den Gleichungen. Eine Ähnlichkeit von 1 ergibt sich für identische Repräsentationen also $A = B$. Dann gilt

$$\begin{aligned} \text{informationsgehalt}(A) &= \\ \text{informationsgehalt}(B) &= \\ \text{informationsgehalt}(\text{gemInfo}(A,B)) & \end{aligned}$$

und es ergibt sich durch Gleichung 4.1.1 und 4.1.2 für die Ähnlichkeit von *A* und *B* ein Wert von 1.

Mit Gleichung 4.1.2 ist es möglich, mehr Fälle der Ähnlichkeit zu unterscheiden, da sich mit Gleichung 4.1.1 für ein bestimmtes Konzept *A* und verschiedene Konzepte *B* der gleiche Wert ergibt. Dies soll an Abbildung 4.2 verdeutlicht werden. Das Konzept *SV* in Abbildung ist der spezifischste Vater der Konzepte. Es entspricht somit der Repräsentation, die gemeinsame Information aus den Konzepten *A*₁, *A*₂, *B*₁ und *B*₂ enthält. Angenommen es soll die Ähnlichkeit von Konzept *A*₂ und den Konzepten *B*₁ und *B*₂ bestimmt werden. Für Gleichung 4.1.2 ergibt sich eine Ähnlichkeit von $\frac{2}{6}$ für *A*₂ und *B*₂ und eine Ähnlichkeit von $\frac{2}{5}$ für *A*₂ und *B*₁. Dieser Unterschied ist mit Gleichung 4.1.1 nicht bestimmbar. Mit Gleichung 4.1.1 ergibt sich nämlich in beiden Fällen eine Ähnlichkeit von $\frac{1}{3}$. Deshalb wird im folgenden Abschnitt 4.2 Gleichung 4.1.2 verwendet.

Angenommen, alle nur möglichen Repräsentationen bilden eine Generalisierungshierarchie, die ähnlich einer Taxonomie aufgebaut ist wie zum

Beispiel in Abbildung 4.2. In der Hierarchie drückt die Subsumptionsrelation also aus, dass die subsumierte Repräsentation mehr oder spezifischere Information enthält als die sie subsumierende Repräsentation. Dann ist die Repräsentation mit der größten gemeinsamen Menge an geteilten Informationen durch die Repräsentation bestimmt, die dem spezifischsten Vater einer Taxonomie entspricht. Folglich lassen sich Gleichungen 4.1.1 und 4.1.2 direkt für Konzepte einer Taxonomie verwenden.

4.2 Ähnlichkeit von Konzepten

Um die Ähnlichkeit von CRIL-Knoten und CRIL-Graphen bestimmen zu können, muss es möglich sein die Ähnlichkeit von Konzepten zu bestimmen. Deshalb wird in diesem Abschnitt ein Ähnlichkeitsmaß für Konzepte einer Taxonomie vorgestellt. Ein Ähnlichkeitsmaß von Konzepten einer Taxonomie soll bestimmte Eigenschaften haben. Diese Eigenschaften betreffen die Verhältnisse der Tiefe der Konzepte zu der Tiefe ihres spezifischsten Vaters.

Eigenschaften eines Ähnlichkeitsmaßes von Konzepten einer Taxonomie:

1. Die Ähnlichkeit von Konzepten ist umso größer, je größer die Tiefe des spezifischsten Vaters ist.
2. Die Ähnlichkeit von Konzepten ist umso kleiner, je größer die Differenz der Tiefe zwischen ihnen und dem spezifischsten Vater ist.
3. Die Ähnlichkeit von Vater und Sohn ist größer als die Ähnlichkeit von Brüdern.

Eigenschaft 1 ist wünschenswert, da in Taxonomien sehr hoch gelegene Konzepte normalerweise für sehr generelle Konzepte stehen, deren Ähnlichkeit untereinander im Allgemeinen geringer ist als die zwischen Konzepten sehr großer Tiefe. So ist es zum Beispiel denkbar, dass in der ersten Hierarchiestufe Konzepte wie Lebewesen und tote Materie repräsentiert werden. Die Ähnlichkeit der beiden Konzepte ist gering. Geht man weiter hinab in der Taxonomie, so kann zum Beispiel das Konzept Affe die Konzepte Schimpanse und Bonobo²⁷ subsumieren. Die Ähnlichkeit zwischen einem Schimpansen und einem Bonobo ist sehr viel größer als die Ähnlichkeit zwischen einem Lebewesen und einem Klumpen Lehm.

Eigenschaft 2 ist wünschenswert, da sie die bereits in Abschnitt 3.2.1 erläuterte geringe Ähnlichkeit weit voneinander entfernter Konzepte zum Ausdruck bringt.

Eigenschaft 3. muss nicht unbedingt wünschenswert sein. So kann es Taxonomien geben, in denen die Ähnlichkeit zwischen Vater und Sohn kleiner als die zwischen Brüdern sein soll. Das oben angeben Beispiel bezüglich der

²⁷ Der Bonobo ist eine dem Schimpansen verwandte Art der Affen.

beiden Affenarten verdeutlicht dies. Im Beispiel des Vergleichs zwischen dem Schimpansen und dem Bonobo sollte die Ähnlichkeit zwischen Schimpansen und Bonobos unter Umständen größer sein als die Ähnlichkeit zwischen einer der beiden Arten und dem Konzept des Affen. Das führt zu folgender alternativer Eigenschaft 3a.

- 3a. Die Ähnlichkeit zwischen Vater und Sohn ist kleiner als die Ähnlichkeit zwischen Brüdern.

Die Verwendung von Gleichung 4.1.2 für die Definition eines Ähnlichkeitsmaßes führt dazu, dass die Eigenschaften 1-3 erfüllt sind. Um die Eigenschaften 1, 2 und 3a zu erfüllen, muss ein anderer Ansatz gewählt werden. Deshalb wird im Anschluss an die Definition der Ähnlichkeit mit Gleichung 4.1.2 ein Ansatz von Zhong, Zhu, Li und Yu aus [ZZL02] vorgestellt, der die Eigenschaften 1, 2 und 3a hat.

Um nun die Gleichung 4.1.2 für die Bestimmung der Ähnlichkeit von Konzepten zu verwenden, lässt sich die Tiefe der Konzepte in einer Taxonomie als Funktion des Informationsgehaltes benutzen. Wie in Abschnitt 3.2.1 bereits deutlich geworden ist, bedeutet eine große Hierarchietiefe, dass das Konzept sehr spezifisch ist und folglich einen hohen Informationsgehalt hat. So enthält zum Beispiel das Konzept Mensch mehr Information als das Konzept Lebewesen, nämlich zum Beispiel Informationen über ungefähre Größe, ungefähres Aussehen und Ähnliches. Das Konzept mit der größten Menge gemeinsamer Information zweier Konzepte ist ihr spezifischster Vater.

Dann lässt sich also in den Gleichungen 4.1.2 die Funktion informationsgehalt durch die Funktion tiefe_T und die Funktion gemInfo durch die Funktion spezVater ersetzen. Das führt zu folgender Definition, in der die Ähnlichkeit zweier Konzepte c_1 und c_2 bezüglich einer Taxonomie T bestimmt wird.

Definition 4.2.1: *Das Maß der Ähnlichkeit zwischen zwei Konzepten c_1 und c_2 einer Taxonomie T ist folgendermaßen definiert:*

$$\text{ähnlichkeit}_T(c_1, c_2) = \frac{2 \text{tiefe}_T(\text{spezVater}_T(c_1, c_2))}{\text{tiefe}_T(c_1) + \text{tiefe}_T(c_2)}$$

Dass die Ähnlichkeitsfunktion aus Definition 4.2.1 die Eigenschaften 1 bis 3 hat, ergibt sich direkt aus der Definition.

Um nun zu erreichen, dass die Ähnlichkeit zwischen Brüdern größer als die Ähnlichkeit zwischen Vater und Sohn ist (Eigenschaft 3a), bietet sich ein anderer Ansatz für die Bestimmung der Ähnlichkeit zweier Konzepte an. Die Ähnlichkeit zweier Konzepte lässt sich auch über ihre konzeptuelle Unterschiedlichkeit²⁸ bestimmen. Bildet man das Maß für die

²⁸ Im Folgenden werde ich anstatt von der konzeptuellen Unterschiedlichkeit nur von der Unterschiedlichkeit sprechen.

Unterschiedlichkeit auf einen Wertebereich zwischen 0 und 1 ab, so lässt sich die Ähnlichkeit durch Subtraktion der Unterschiedlichkeit von 1 berechnen.

Definition 4.2.2: Sei $\text{unterschiedlichkeit}_T(c_1, c_2)$ ein Maß für die Unterschiedlichkeit zweier Konzepte c_1 und c_2 einer Taxonomie T und es gelte

$$0 \leq \text{unterschiedlichkeit}_T(c_1, c_2) \leq 1.$$

Dann ist das Maß der Ähnlichkeit zweier Konzepte c_1 und c_2 folgendermaßen definiert:

$$\text{ähnlichkeit}_T(c_1, c_2) = 1 - \text{unterschiedlichkeit}_T(c_1, c_2)$$

Für die Konzepte einer Taxonomie besteht die Möglichkeit, ein Maß für die Unterschiedlichkeit zu bestimmen, darin, die Entfernung zwischen den Konzepten in der Taxonomie zu ermitteln. Je größer die Entfernung zwischen den beiden Konzepten, desto größer die Unterschiedlichkeit. Dies bringt zum Ausdruck, dass Konzepte, die in der Taxonomie weit voneinander entfernt liegen, als wenig ähnlich angesehen werden (vgl. oben). Ein intuitiver Ansatz für eine solche Bestimmung eines Maßes für die Unterschiedlichkeit ist durch die folgende Definition 4.2.3 gegeben.

Definition 4.2.3: Das Maß für die Unterschiedlichkeit zweier Konzepte c_1 und c_2 einer Taxonomie T ist folgendermaßen definiert:

$$\text{unterschiedlichkeit}_T(c_1, c_2) = \frac{\text{entfernung}_T(c_1, c_2, \text{spezVater}_T(c_1, c_2))}{\max\text{PL}_T}$$

Definition 4.2.3 hat nun die gewünschte Eigenschaft 3a bezüglich des Verhältnisses der Ähnlichkeit zwischen Brüdern und der Ähnlichkeit zwischen Vater und Sohn, allerdings nicht die Eigenschaften 1 und 2.

Einen Vorschlag für ein Ähnlichkeitsmaß, das die Eigenschaften 1, 2, und 3a hat, machen Zhung et al. in [ZZL02]. Zhung et al. bestimmen ebenfalls zunächst die Unterschiedlichkeit. Dabei wird jedem Konzept ein von Zhung et al. so genannter Markstein, also eine entscheidende Kennzeichnung, in Abhängigkeit von der Tiefe des Konzepts in der Taxonomie zugeordnet. Man beachte, dass der Markstein mit zunehmender Tiefe abnimmt.

Definition 4.2.4: Gegeben eine streng monoton steigende Funktion f über natürlichen Zahlen für die gilt $f(0) = 1$. Dann ist der Markstein eines Konzepts c einer Taxonomie T bezüglich f folgendermaßen definiert:

$$\text{markstein}_{T,f}(c) = \frac{0,5}{f(\text{tiefe}_T(c))}$$

Der Dividend in der Gleichung von Definition 4.2.4 ist 0,5, damit die Funktion, die das Maß der Unterschiedlichkeit bestimmt (vgl. Definition 4.2.5), auf das Intervall $[0,1)$ normiert wird, also den gewünschten Wertebereich hat. Das Maß der Unterschiedlichkeit zwischen zwei Konzepten ist dann nach [ZZL02] folgendermaßen definiert.

Definition 4.2.5: *Gegeben sei eine streng monoton steigende Funktion f über natürlichen Zahlen, für die gilt $f(0) = 1$. Dann ist das Maß der Unterschiedlichkeit von zwei Konzepten c_1 und c_2 einer Taxonomie T bezüglich f folgendermaßen definiert:*

$$\text{unterschiedlichkeit}_{T,f}(c_1, c_2) = \text{unterschiedSV}_{T,f}(c_1, \text{spezVater}_T(c_1, c_2)) + \text{unterschiedSV}_{T,f}(c_2, \text{spezVater}_T(c_1, c_2))$$

mit

$$\text{unterschiedSV}_{T,f}(c, c') = \text{markstein}_{T,f}(c') - \text{markstein}_{T,f}(c)$$

Formt man die Gleichung um, so wird deutlich, dass der oben bereits erwähnte Dividend in der Gleichung von Definition 4.2.4 0,5 sein muss.

$$\text{unterschiedlichkeit}_{T,f}(c_1, c_2) = 2 \text{ markstein}(\text{spezVater}_T(c_1, c_2)) - (\text{markstein}_{T,f}(c_1) + \text{markstein}_{T,f}(c_2))$$

Die maximale Unterschiedlichkeit zweier Konzepte wird erreicht, wenn der spezifischste Vater T ist und die Tiefe der beiden verglichenen Konzepte maximal, also sehr groß, ist. In diesem Fall sollte die Unterschiedlichkeit der Konzepte nahe null oder gleich null sein, was für die Funktion $\text{unterschiedlichkeit}_{T,f}$ offensichtlich zutrifft, wie die Umformung zeigt.

In [ZZL02] wird als Funktion $f(x) = k^x$ verwendet, wobei gelten soll $k > 1$. Dies bewirkt, dass der Informationsgehalt der Konzepte exponentiell mit der Tiefe der Konzepte zunimmt, da die Unterschiedlichkeit exponentiell mit der Entfernung vom spezifischsten Vater abnimmt. Die Frage, in welchen Fällen der Modellierung einer Taxonomie eine solche Bestimmung der Ähnlichkeit sinnvoll ist, soll an dieser Stelle nicht beantwortet werden.

Ob nun ein Maß der Ähnlichkeit mit der Eigenschaft 3 oder der Eigenschaft 3a verwendet werden sollte, hängt m. E. von der Modellierung der Taxonomie ab. Die in diesem Abschnitt verwendeten Beispiele machen deutlich, dass es sowohl sinnvoll sein kann eine Ähnlichkeitsfunktion mit Eigenschaft 3 zu verwenden, als auch eine Ähnlichkeitsfunktion mit Eigenschaft 3a. Welche der Funktionen verwendet wird hängt nicht zuletzt von der Art der Modellierung der Taxonomien ab. Mit Hilfe des Geometrischen Agenten kann untersucht werden, welche Eigenschaft zu einer besseren Navigationsleistung bei einer bestimmten Modellierung der Taxonomie führt.

Generell sollte m. E. zusätzlich noch folgendes in die Bestimmung der Ähnlichkeit mit einfließen:

- Ist spezifischste Vater der verglichenen Konzepte T , so sollte die Ähnlichkeit als sehr gering angenommen werden.²⁹
- Stehen die verglichenen Konzepte in der Exklusionsrelation zueinander, so sollte die Ähnlichkeit der Konzepte als sehr gering angenommen werden.

Aus diesen Annahmen lässt sich nun eine Modifikation der Ähnlichkeitsfunktion vornehmen, die in der folgenden Definition 4.2.6 festgehalten ist.

Definition 4.2.6: Seien c_1 und c_2 zwei Konzepte aus einer Taxonomie $T = (S, \leq, \leftrightarrow)$ und $\text{ähnlichkeit}_T(c_1, c_2)$ entweder durch Definition 4.2.1 oder durch Definition 4.2.2 bestimmt. Dann ist das Maß der Ähnlichkeit zwischen den beiden Konzepten durch die Funktion $\text{ähnlichkeit}'_T(c_1, c_2)$ bestimmt.

$$\text{ähnlichkeit}'_T(c_1, c_2) = \begin{cases} 0 & , \text{wenn } (\text{spezVater}(c_1, c_2) = T) \vee c_1 \leftrightarrow c_2 \\ \text{ähnlichkeit}_T(c_1, c_2) & , \text{sonst} \end{cases}$$

Die in diesem Abschnitt vorgeschlagenen Ähnlichkeitsmaße dienen dem Vergleich von Konzepten, die zum einen Wissen einer natürlichsprachlichen Beschreibung (der Wegbeschreibung) repräsentieren und zum anderen Wissen des Agenten repräsentieren, das aus der Perzeption stammt. Ist das wahrgenommene Konzept (P-Konzept) genereller als das Konzept der Instruktion (I-Konzept), so kann es sein, dass das P-Konzept eben nicht koreferent zum I-Konzept ist oder dass der Instrukteur eine falsche Wegbeschreibung gegeben hat. In diesem Fall kann die Ähnlichkeit der Konzepte wie oben beschrieben bestimmt werden.

Im umgekehrten Fall, also dass das I-Konzept genereller ist als das P-Konzept, lässt sich eine solche Schlussfolgerung nicht zwingend treffen. Der Instrukteur könnte bewusst ein allgemeines Konzept verwendet haben, beispielsweise da er sich nicht sicher war oder da ihm das allgemeine Konzept als ausreichend für die Beschreibung des Weges erschien. Ist dies der Fall, sollte für jedes Konzept P der Perzeption, das von dem Konzept I der Instruktion subsumiert wird, die Ähnlichkeit von P und I groß sein.

Folgende Beispiele sollen dies verdeutlichen. Angenommen, in einer Wegbeschreibung wird das Konzept Gebäude verwendet und das Konzept Gebäude subsumiert die Konzepte Haus und Turm. Wird nun während der Navigation (zum entsprechenden Zeitpunkt) ein Haus oder ein Turm wahrgenommen, so ist es gerechtfertigt anzunehmen, dass es sich um das Objekt der Wegbeschreibung handelt. Wird umgekehrt in der

²⁹ Ich nehme im Folgenden eine Ähnlichkeit von 0 an. Da nicht sicher ist, ob diese Annahme generell richtig ist, spreche ich hier allgemein von einer geringen Ähnlichkeit.

Wegbeschreibung von einem Turm oder einem Haus gesprochen und ein Objekt wahrgenommen, das ein nicht genauer bestimmbares Gebäude ist, so ist der Schluss, dass es sich um das Objekt der Wegbeschreibung handelt, nicht unbedingt gerechtfertigt. Zumindest nicht, wenn nicht noch andere Eigenschaften des Objekts wahrgenommen wurden, die doch für eine hohe Ähnlichkeit sprechen.

Diese Abhängigkeit der Ähnlichkeit von der Herkunft der verglichenen Konzepte sollte unter Umständen in die Bestimmung der Ähnlichkeit einfließen. Generell muss dafür bei der Bestimmung der Ähnlichkeit zwischen den Konzepten unterschieden werden. Für den GA könnte das Vorgehen dann folgendermaßen aussehen. Die Ähnlichkeit zwischen einem I-Konzept und einer P-Konzept könnte dann als sehr groß angesehen werden, wenn das P-Konzept von dem I-Konzept subsumiert wird. Ist dies nicht der Fall würde die Ähnlichkeit nach Definition 4.2.6 bestimmt werden. Wie genau die Ähnlichkeit zwischen den Konzepten in solchen Fällen variiert, bedarf m. E. einer genaueren psychologischen und linguistischen Analyse und kann nicht Gegenstand dieser Arbeit sein. Für die im Folgenden vorgestellten Bestimmungen der Ähnlichkeiten von Knoten und Graphen wird jeweils ein Vorschlag gemacht, der die Herkunft der Repräsentation berücksichtigt.

4.3 Ähnlichkeit von Knoten

Die in einem CRIL-Knoten enthaltene Information ist bestimmt durch die Mengen der Sorten, Eigenschaften und Namen. Wie die Ähnlichkeit von Konzepten wie Sorten und Eigenschaften bestimmt werden kann, wurde im letzten Abschnitt gezeigt. Um die Ähnlichkeit zweier CRIL-Knoten zu bestimmen, bietet es sich an, die Ähnlichkeiten der in ihnen enthaltenen Konzepte zu verwenden. Dafür soll je eine Sorte (bzw. Eigenschaft) aus einem der Knoten mit maximal einer Sorte (bzw. Eigenschaft) aus dem anderen Knoten verglichen werden. Dann wird für jedes so gebildete Paar von Konzepten die Ähnlichkeit bestimmt. Die Summe der so bestimmten Ähnlichkeiten ist bereits ein Maß für die Ähnlichkeit der verglichenen Knoten. Um den gewünschten Wertebereich zwischen 0 und 1 zu erhalten, lässt sich die Summe noch normieren. Die Normierung kann auf verschiedene Weisen erfolgen. Für den Vergleich eines P-Knotens mit einem I-Knoten sollten die in dem I-Knoten enthaltenen Konzepte für die Normierung verwendet werden. Wenn ein I-Knoten zum Beispiel nur ein Konzept enthält wie etwa die Sorte Gebäude, dann sollten alle P-Knoten ähnlich zu dem I-Knoten sein, die mindestens die Sorte Gebäude oder eine ähnliche Sorte enthalten. Wenn andererseits ein I-Knoten außer der Sorte Haus noch mehrere andere Konzepte enthält, so sollten nur P-Knoten ähnlich zu dem I-Knoten sein, die entsprechend viele Konzepte enthalten, die ähnlich zu den Konzepten des I-Knoten sind. Dies kann erreicht werden, indem die oben beschriebene Summe der Ähnlichkeiten auf die Anzahl der Konzepte des I-Knoten normiert wird.

Wie dieses Ähnlichkeitsmaß bestimmt werden kann, wird im Folgenden gezeigt. Es werden ferner zwei weitere Normierungen vorgeschlagen, die andere Eigenschaften aufweisen.

Um die eben beschriebenen Konzeptpaare von CRIL-Knoten zu bestimmen, soll zunächst eine Verknüpfung zweier CRIL-Graphen definiert werden. Eine Verknüpfung besteht aus zwei Mengen von Paaren, eine Menge von Paaren bestehend aus Sorten und eine Menge von Paaren bestehend aus Eigenschaften.

Definition 4.3.1: Für zwei CRIL-Knoten $ck_1 = (SOR_1, EIG_1, NAM_1)$ und $ck_2 = (SOR_2, EIG_2, NAM_2)$ ist eine Verknüpfung $match(ck_1, ck_2) = (SV, EV)$ ein Paar, bestehend aus der Menge $SV \subseteq SOR_1 \times SOR_2$ der Sortenverknüpfungen und der Menge $EV \subseteq EIG_1 \times EIG_2$ der Eigenschaftsverknüpfungen.

Für eine Verknüpfung $match(ck_1, ck_2) = (SV, EV)$ sind folgende Funktionen definiert:

$$\begin{aligned} S\mathcal{V}(match(ck_1, ck_2)) &= SV \\ E\mathcal{V}(match(ck_1, ck_2)) &= EV \end{aligned}$$

Für eine gegebene Verknüpfung kann nun für jedes in ihr enthaltene Paar die Ähnlichkeit der Elemente des Paares bestimmt werden. Die Summe der so bestimmten Ähnlichkeiten sei durch die im folgenden definierte Funktion Ähnlichkeitsgehalt bestimmt.

Definition 4.3.2: Gegeben eine Taxonomie $T = (S, \leq, \leftrightarrow)$ und zwei CRIL-Knoten $ck_1 = (SOR_1, EIG_1, NAM_1)$ und $ck_2 = (SOR_2, EIG_2, NAM_2)$, für die gilt $SOR_1 \cup SOR_2 \cup EIG_1 \cup EIG_2 \subseteq S$. Der Ähnlichkeitsgehalt einer Verknüpfung $match(ck_1, ck_2)$ bezüglich T ist folgendermaßen bestimmt:

$$\text{Ähnlichkeitsgehalt}_T(match(ck_1, ck_2)) = \sum_{(s_1, s_2) \in S\mathcal{V}(match(ck_1, ck_2))} \text{Ähnlichkeit}'_T(s_1, s_2) + \sum_{(e_1, e_2) \in E\mathcal{V}(match(ck_1, ck_2))} \text{Ähnlichkeit}'_T(e_1, e_2)$$

Mit dem durch Definition 4.3.2 bestimmten Wert des Ähnlichkeitsgehalts lässt sich nun ein Ähnlichkeitsmaß für CRIL-Knoten bezüglich einer Verknüpfung bestimmen. In Definition 4.3.3 wird nun, wie eingangs beschrieben, der Ähnlichkeitsgehalt auf die Anzahl der im I-Knoten enthaltenen Konzepte normiert.

Definition 4.3.3: Gegeben die Taxonomie $T = (S, \leq, \leftrightarrow)$ der Attribute, einen I-Knoten $ck_I = (SOR_I, EIG_I, NAM_I)$ und einem P-Knoten $ck_P = (SOR_P, EIG_P, NAM_P)$, für die gilt $SOR_I \cup SOR_P \cup EIG_I \cup EIG_P \subseteq S$. Das Maß der Ähnlichkeit von ck_I und ck_P bezüglich der Verknüpfung $vk = match(ck_I, ck_P)$ und T ist folgendermaßen bestimmt.

$$\text{Ähnlichkeit}_T(ck_I, ck_P, vk) = \frac{\text{Ähnlichkeitsgehalt}_T(vk)}{|SOR_I| + |EIG_I|}$$

Alternativ zu Definition 4.3.3 lässt sich die Ähnlichkeit auch analog zu Gleichung 4.1.2 bestimmen (4.3.4 a) oder ähnlich zu Gleichung 4.1.1 (4.3.4 b). Der Informationsgehalt der Knoten ist dann die Anzahl der in ihnen enthaltenen Konzepte. Der Informationsgehalt der Repräsentation, die gemeinsame Informationen aus den verglichenen Knoten enthält, ist der Ähnlichkeitsgehalt nach Definition 4.3.2. Dass die Verknüpfung als eine Repräsentation angesehen werden kann, die gemeinsame Informationen aus den verknüpften Knoten enthält, ergibt sich daraus, dass jedes Paar der Verknüpfung einen spezifischsten Vater repräsentiert.

Definition 4.3.3: Gegeben eine Taxonomie $T = (S, \leq, \leftrightarrow)$ und zwei CRIL-Knoten $ck_1 = (SOR_1, EIG_1, NAM_1)$ und $ck_2 = (SOR_2, EIG_2, NAM_2)$, für die gilt $SOR_1 \cup SOR_2 \cup EIG_1 \cup EIG_2 \subseteq S$. Das Maß der Ähnlichkeit von ck_1 und ck_2 bezüglich der Verknüpfung $vk = match(ck_1, ck_2)$ und T ist folgendermaßen bestimmt.

(4.3.4 a)

$$\text{ähnlichkeit}_T(ck_1, ck_2, vk) = \frac{2 \text{ähnlichkeitsgehalt}_T(vk)}{|\text{SOR}_1| + |\text{EIG}_1| + |\text{SOR}_2| + |\text{EIG}_2|}$$

(4.3.4 b)

$$\text{ähnlichkeit}_T(ck_1, ck_2, vk) = \frac{\text{ähnlichkeitsgehalt}_T(vk)}{\min((|\text{SOR}_1| + |\text{EIG}_1|), (|\text{SOR}_2| + |\text{EIG}_2|))}$$

Die durch 4.3.4a bestimmte Ähnlichkeit hat die Eigenschaft, für Knoten, die sich in der Anzahl der enthaltenen Konzepte stark unterscheiden, generell klein zu sein. Die durch 4.3.4b bestimmte Ähnlichkeit hingegen, kann auch bei großen Unterschieden in der Konzeptzahl groß sein.

Eine Verknüpfung nach Definition 4.3.1 lässt noch Paare zu, in denen zum Beispiel die ersten Elemente gleich sind, also etwa die Paare (x_1, y_1) und (x_2, y_2) , in denen $x_1 = x_2$ ist. Solche Paare sollen aber ausgeschlossen sein (vgl. oben). Außerdem kann für eine solche Verknüpfung die Ähnlichkeit nach Definition 4.3.3 (oder 4.3.4) größer 1 sein. Es muss also die Verknüpfung gefunden werden,

- die konsistent ist, also nur Paare enthält, deren Elemente entweder aus verschiedenen Knoten stammen oder paarweise verschieden sind, und
- für die es keine konsistente Verknüpfung gibt, die eine größere Ähnlichkeit der verknüpften Knoten nach Definition 4.3.3 ergibt.

Um die gesuchte Verknüpfung zu finden, muss es möglich sein zu bestimmen, ob eine gegebene Verknüpfung konsistent ist.

Definition 4.3.4: Gegeben zwei CRIL-Knoten ck_1 und ck_2 . Eine Verknüpfung von ck_1 und ck_2 ist konsistent, wenn die in ihren Paaren enthaltenen Konzepte eines der Knoten paarweise verschieden sind.

$$\begin{aligned} \text{kons}(\text{match}(ck_1, ck_2)) \stackrel{\text{def}}{\Leftrightarrow} & \forall (s_1, s_2), (s_3, s_4) \in \mathcal{SV}(\text{match}(ck_1, ck_2)): \\ & [(s_1, s_2) = (s_3, s_4)] \vee [(s_1 \neq s_3) \wedge (s_2 \neq s_4)] \\ & \forall (e_1, e_2), (e_3, e_4) \in \mathcal{EV}(\text{match}(ck_1, ck_2)): \\ & [(e_1, e_2) = (e_3, e_4)] \vee [(e_1 \neq e_3) \wedge (e_2 \neq e_4)] \end{aligned}$$

Nun kann die gesuchte Verknüpfung gefunden werden, indem bezüglich aller konsistenten Verknüpfungen der verglichenen Knoten die Ähnlichkeit nach Definition 4.3.3 oder 4.3.4 bestimmt wird. Die Verknüpfung mit der größten Ähnlichkeit ist die gesuchte Verknüpfung. Die bezüglich dieser Verknüpfung bestimmte Ähnlichkeit soll die Ähnlichkeit der Knoten bestimmen. Die Ähnlichkeit von zwei CRIL-Knoten kann dann folgendermaßen bestimmt werden.

Definition 4.3.5: Gegeben eine Taxonomie $T = (S, \leq, \leftrightarrow)$ und zwei CRIL-Knoten $ck_1 = (SOR_1, EIG_1, NAM_1)$ und $ck_2 = (SOR_2, EIG_2, NAM_2)$, für die gilt $SOR_1 \cup SOR_2 \cup EIG_1 \cup EIG_2 \subseteq S$. Für ck_1 und ck_2 und deren Verknüpfung $vk = \text{match}(ck_1, ck_2)$, für die folgendes gilt

$$\begin{aligned} \forall vk' = \text{match}(ck_1, ck_2): \\ \text{ähnlichkeit}_T(ck_1, ck_2, vk) \geq \text{ähnlichkeit}_T(ck_1, ck_2, vk') \end{aligned}$$

ist die Ähnlichkeit von ck_1 und ck_2 bestimmt durch

$$\text{ähnlichkeit}_T(ck_1, ck_2) = \text{ähnlichkeit}_T(ck_1, ck_2, vk).$$

Bei der Bestimmung der Ähnlichkeit von zwei CRIL-Knoten sollen noch folgende Punkte berücksichtigt werden.

1. Wenn es Sorten aus den beiden CRIL-Knoten gibt, die in der Exklusionsrelation zueinander stehen, so führt dies zu einer besonders niedrigen Ähnlichkeit.
2. Übereinstimmungen zwischen anderen Eigenschaften, wie zum Beispiel Namen, aber auch bestimmten Sorten und Eigenschaften führen zu einer besonders hohen Ähnlichkeit, solange 1 nicht verletzt ist.

Um Punkt 1 zu garantieren, wird die Funktion komp über zwei CRIL-Knoten definiert. Mit ihr kann geprüft werden, ob alle Sorten der CRIL-Knoten

kompatibel zueinander sind, also nicht in der Exklusionsrelation zueinander stehen.

Definition 4.3.6: Gegeben seien zwei CRIL-Graphen $G1$ und $G2$ und die zu beiden Graphen gehörenden Taxonomien $T_A = (S_A, \leq_A, \leftrightarrow_A)$ der Attribute und $T_R = (S_R, \leq_R, \leftrightarrow_R)$ der Relationen. Zwei CRIL-Knoten $ck_1 = (SOR_1, EIG_1, NAM_1)$ und $ck_2 = (SOR_2, EIG_2, NAM_2)$ aus $G1$ oder $G2$ sind kompatibel zueinander, wenn es keine Sorte aus SOR_1 gibt, die in der Exklusionsrelation zu den Sorten aus SOR_2 steht.

$$\text{komp}(ck_1, ck_2) \stackrel{\text{def}}{\Leftrightarrow} \neg \exists s_1, s_2: s_1 \in SOR_1 \wedge s_2 \in SOR_2 \wedge s_1 \leftrightarrow_A s_2$$

Um Punkt 2 bei der Bestimmung der Ähnlichkeit zweier CRIL-Knoten zu prüfen, ist es nötig, bei einem Vergleich eine Menge von Sorten und Eigenschaften zu definieren, die für eine hohe Ähnlichkeit sprechen. Diese Sorten und Eigenschaften sollen durch die Menge IdentAtt bestimmt sein.

Für einen Vergleich von CRIL-Knoten aus CRIL-Graphen, zu denen eine Taxonomie $T_A = (S_A, \leq_A, \leftrightarrow_A)$ gehört, ist die Menge der Übereinstimmungsattribute IdentAtt Teilmenge von S_A .

$$\text{IdentAtt} \subset S_A$$

Prinzipiell wäre es auch denkbar, dass die Menge IdentAtt auch Relationen enthält. Die Einbeziehung der Relationen in das Ähnlichkeitsmaß erfolgt allerdings bei der Bestimmung der Ähnlichkeit zwischen Graphen und ist deshalb an dieser Stelle nicht nötig. Zwei gegebenen CRIL-Knoten sollen nun eine hohe Übereinstimmung aufweisen, wenn ein oder mehrere Elemente aus IdentAtt in den Sorten- oder Eigenschaftsmengen beider Knoten enthalten ist.

Definition 4.3.7:

$$\text{ident}(ck_1, ck_2) \stackrel{\text{def}}{\Leftrightarrow} [\exists a \in \text{IdentAtt}: a \in SOR_1 \wedge a \in SOR_2] \vee [\exists n_1, n_2: n_1 \in NAM_1 \wedge n_2 \in NAM_2 \wedge n_1 = n_2].$$

Nun lässt sich die Ähnlichkeitsfunktion über zwei CRIL-Knoten ähnlich wie in Definition 4.2.6 definieren, wobei die Punkte 1 und 2 berücksichtigt werden.

Definition 4.3.8:

$$\text{ähnlichkeit}'_T(ck_1, ck_2) = \begin{cases} 0 & , \text{wenn } \neg \text{komp}(ck_1, ck_2) \\ 1 & , \text{wenn } \text{komp}(ck_1, ck_2) \wedge \text{ident}(ck_1, ck_2) \\ \text{ähnlichkeit}(ck_1, ck_2) & , \text{sonst} \end{cases}$$

4.4 Ähnlichkeit von Graphen

Die Ähnlichkeit zweier CRIL-Graphen kann nun in ganz ähnlicher Weise bestimmt werden. Dafür sei zunächst die Verknüpfung von zwei CRIL-Graphen definiert. Sie besteht aus einer Menge von Knotenpaaren und einer Menge von Kantenpaaren. Über die Kantenpaare soll die Ähnlichkeit der durch die verknüpften Kanten repräsentierten Konzepte bestimmt werden können. Dass bedeutet, dass ein Kantenpaar nur dann sinnvoll ist, wenn der spezifischste Vater der durch die Kanten repräsentierten Konzepte nicht T ist. Dies lässt sich feststellen, indem der Kantenoperator auf die Kante angewendet wird und die so bestimmten Konzepte verglichen werden. Ferner soll es für eine Kante eines verglichenen Graphen nur maximal eine Verknüpfung mit einer Kante des anderen Graphen geben. Dies kann in die Definition der Verknüpfung mit aufgenommen werden. Dass diese Konsistenzforderung bezüglich der Kanten des Graphen mit in die Definition der Verknüpfung aufgenommen wird, hat den Grund, dass in Abschnitt 4.5 nur solche Verknüpfungen verwendet werden sollen.

Definition 4.4.1: Für zwei CRIL-Graphen $G1 = (V_1, E_1, T_{A1}, T_{R1}, op_1)$ und $G2 = (V_2, E_2, T_{A2}, T_{R2}, op_2)$ ist eine Verknüpfung $match(G1, G2) = (VV, EV)$ ein Paar, bestehend aus der Menge $VV \subseteq V_1 \times V_2$ der Knotenverknüpfungen und der Menge $EV \subseteq E_1 \times E_2$ der Kantenverknüpfungen, für die gilt

$$\forall (e_1, e_2) \in EV: e_1 = (ck_1, ck'_1) \wedge e_2 = (ck_2, ck'_2) \Rightarrow \\ [\exists (ck_1, ck_2), (ck'_1, ck'_2) \in VV \wedge spezVater(op_1(r_1), op_2(r_2)) \neq T]$$

Für eine Verknüpfung $match(G1, G2) = (VV, EV)$ sind folgende Funktionen definiert:

$$\mathcal{V}\mathcal{V}(match(G1, G2)) = VV \\ \mathcal{E}\mathcal{V}(match(G1, G2)) = EV$$

Nun kann analog zu Definition 4.3.2 der Ähnlichkeitsgehalt einer Verknüpfung von CRIL-Graphen bestimmt werden.

Definition 4.4.2: Gegeben seien die Taxonomien T_A und T_R . Für zwei CRIL-Graphen $G1 = (V_1, E_1, T_A, T_R, op_1)$ und $G2 = (V_2, E_2, T_A, T_R, op_2)$ ist der Ähnlichkeitsgehalt einer Verknüpfung $match(G1, G2)$ bestimmt durch

$$\text{ähnlichkeitsgehalt}(match(G1, G2)) = \sum_{(ck_1, ck_2) \in \mathcal{V}\mathcal{V}(match(G1, G2))} \text{ähnlichkeit}_{T_A}'(ck_1, ck_2) \\ + \sum_{(r_1, r_2) \in \mathcal{E}\mathcal{V}(match(G1, G2))} \text{ähnlichkeit}_{T_R}'(op_1(r_1), op_2(r_2))$$

Mit dem Ähnlichkeitsgehalt einer Verknüpfung kann wiederum die Ähnlichkeit zweier CRIL-Graphen bezüglich der Verknüpfung bestimmt werden. Zunächst wird wieder die Ähnlichkeit für einen Vergleich eines I-Graphen mit einem P-Graphen definiert. Dabei wird auf der Ähnlichkeitsgehalt auf die Anzahl der im I-Graphen enthaltenen Konzepte normiert.

Definition 4.4.3: Gegeben seien die Taxonomien T_A und T_R . Das Maß der Ähnlichkeit von einem I-Graphen $G_I = (V_I, E_I, T_A, T_R, op_I)$ und einem P-Graphen $G_P = (V_P, E_P, T_A, T_R, op_P)$ bezüglich einer Verknüpfung $match(G1, G2)$ ist bestimmt durch

$$\text{ähnlichkeit}(G1, G2, match(G1, G2)) = \frac{\text{ähnlichkeitsgehalt}(match(G1, G2))}{|V_I| + |E_I|}$$

Alternativ dazu lässt sich wieder eine zu Definition 4.3.4a analoge Definition der Ähnlichkeit verwenden. Für die Bestimmung der Ähnlichkeit von I-Graphen und P-Graphen sollte aber Definition 4.4.3 verwendet werden.

Definition 4.4.4: Gegeben seien die Taxonomien T_A und T_R . Für zwei CRIL-Graphen $G1 = (V_1, E_1, T_A, T_R, op_1)$ und $G2 = (V_2, E_2, T_A, T_R, op_2)$ ist das Maß der Ähnlichkeit bezüglich einer Verknüpfung $match(G1, G2)$ bestimmt durch

$$\text{ähnlichkeit}(G1, G2, match(G1, G2)) = \frac{2 \text{ähnlichkeitsgehalt}(match(G1, G2))}{|V_1| + |V_2| + |E_1| + |E_2|}$$

Es muss wieder geprüft werden können, ob eine Verknüpfung konsistent ist, also ob die Elemente der Knotenpaare und der Kantenpaare paarweise verschieden in den verglichenen Graphen sind.

Definition 4.4.5: Gegeben zwei CRIL-Graphen $G1$ und $G2$. Die Funktion $kons$ über eine Verknüpfung von $G1$ und $G2$ ist definiert durch

$$\text{kons}(match(G1, G2)) \stackrel{\text{def}}{\Leftrightarrow} \forall (ck_1, ck_2), (ck_3, ck_4) \in \mathcal{V}\mathcal{V}(match(G1, G2)): \\ [(ck_1, s_2) = (ck_3, ck_4)] \vee [(ck_1 \neq ck_3) \wedge (ck_2 \neq ck_4)]$$

Es fällt auf, dass die Konsistenz der Kantenpaare in Definition 4.4.5 nicht geprüft wird. Diese Prüfung ist nicht nötig, da sich die Konsistenz der Kantenpaare aus Definition 4.4.1 ergibt.

Bei der Bestimmung der Ähnlichkeit von zwei Graphen ließe sich ebenfalls berücksichtigen, dass das Vorhandensein einer bestimmten durch ein Kantenpaar repräsentierten Relation in der Verknüpfung die Ähnlichkeit der verglichenen Graphen erhöht. Dafür könnte eine Menge $ParRel$ definiert werden, in der die entsprechenden Konzepte der Relationen enthalten sind. Entspricht das Konzept aus $ParRel$ einem spezifischsten Vater, der für die

Bestimmung der Ähnlichkeit bezüglich eines Kantenpaares verwendet wurde, so kann dies, in einer hier nicht näher spezifizierten Weise, das Maß der Ähnlichkeit zwischen den Graphen erhöhen. Es sei angemerkt, dass die benötigte Definition nicht die Form von Definition 4.2.6 haben darf, da das Vorhandensein einer einzelnen Relation nicht zu einer Ähnlichkeit von 1 führen darf.

Nun noch einige abschließende Bemerkungen bezüglich der Ähnlichkeit zwischen graphenbasierten Repräsentationen. Nach [PoC95] gibt es für Graphen drei unterschiedliche Aspekte von Ähnlichkeit. Diese sind wie folgt.

- **Oberflächliche Ähnlichkeit**
Basiert auf der Ähnlichkeit von Objekten und Attributen. Relationen oder Muster werden nicht berücksichtigt.
- **Strukturelle Ähnlichkeit**
Basiert auf Ähnlichkeiten bezüglich Dingen wie z. B. Anzahl von Kanten, Knotengrad oder Pfadlängen.
- **Thematische Ähnlichkeit**
Bedeutung basiert auf dem Vorhandensein von Relationen zwischen Knoten und bestimmten Mustern von Konzepten und Relationen.

Die oberflächliche Ähnlichkeit beinhaltet nur die Ähnlichkeit zwischen Konzepten oder, im Fall von CRIL-Graphen, Knoten und berücksichtigt damit nur einen sehr geringen Teil der enthaltenen Informationen. Die strukturelle Ähnlichkeit, die Strukturen wie Kantenlänge oder Knotengrade verwendet, ist für einen Vergleich von Graphen, die natürlichsprachliche Konzepte repräsentieren, eher ungeeignet, obwohl zum Beispiel in [MGL00b] der Knotengrad in die Berechnung der Ähnlichkeit von konzeptuellen Graphen mit einfließt.

Bei der vorliegenden Definition von Ähnlichkeit handelt es sich m. E. und im Gegensatz zu Zhung et al. in [ZZL02], nach denen es sich um eine rein thematische Ähnlichkeit handelt, um eine Art oberflächlich-thematische Ähnlichkeit. Zwar werden Relationen von Objekten berücksichtigt, aber in der Beschreibung von thematischer Ähnlichkeit in [PoC95] wird zusätzlich von domänenspezifischen Mustern oder Mustern, die aus skriptähnlichem Wissen entstehen, gesprochen.

Solche Muster sind m. E. zum Beispiel Graphen, die in WB enthalten sind und bestimmtes Wissen repräsentieren. Diese Graphen könnten bei der Bestimmung der Ähnlichkeit zweier Graphen mit einbezogen werden, indem zum Beispiel nach dem Vorhandensein eines dem Muster entsprechenden Teilgraphen in den beiden verglichenen Graphen oder dem BMG gesucht wird, und sich im Fall des Vorhandenseins ein hohes oder erhöhtes Ähnlichkeitsmaß ergibt. Solche Muster werden bei der vorliegenden Berechnung der Ähnlichkeit nicht berücksichtigt, weshalb die Bezeichnung m. E. unpassend ist.

Eine Möglichkeit, solche Muster in die Bestimmung der Ähnlichkeit mit einfließen zu lassen, wäre zum Beispiel, anstatt der oben erwähnten Menge ParRel eine Menge ParMuster, die ganze Teilgraphen enthält, zu verwenden. Liegt ein bestimmtes Muster in beiden Graphen vor, so könnte die Ähnlichkeit als sehr hoch angenommen werden.

4.5 Berechnung der Ähnlichkeit zweier Graphen

In diesem Abschnitt wird nun beschrieben, wie die Verknüpfung zweier CRIL-Graphen berechnet werden kann, die konsistent ist und die die größte Ähnlichkeit nach Definition 4.4.4 ergibt. Der Algorithmus basiert auf dem von Poole und Campbell in [PoC95] vorgestellten Verfahren, das aus zwei Schritten besteht. Der erste Schritt des Verfahrens besteht im Bilden einer Verknüpfung, in der zum einen das kartesische Produkt der Knotenmengen der verglichenen Knoten enthalten ist, sowie zum anderen das kartesische Produkt der Kanten, soweit dies nach Definition 4.4.1 möglich ist. Im zweiten Schritt werden solange Knoten und Kanten aus der Verknüpfung gelöscht, bis die gesuchte Verknüpfung gefunden ist.

Für den Algorithmus des ersten Schritts wird eine Prozedur VERKNÜPFE benötigt, die auf zwei Knoten oder Kanten eines CRIL-Graphen anwendbar sein muss. Sie gibt genau dann ein Knoten- oder Kantenpaar zurück, wenn die Ähnlichkeit der Argumente größer 0 ist. Sonst gibt VERKNÜPFE nil zurück. Für die Verknüpfung der Kanten wird eine Matrix aus den Knoten der beiden Graphen aufgebaut. Nach der Verknüpfung der Knoten wird das Ergebnis der Prozedur VERKNÜPFE in der Matrix unter den Argumenten der Prozedur eingetragen (vgl. Zeile 6 in Algorithmus 4.5.1). Wird die Prozedur VERKNÜPFE auf zwei Relationen angewendet, wird überprüft ob

1. die Ähnlichkeit der durch die Kanten repräsentierten Konzepte größer 0 ist (der spezifischste Vater der durch die Kanten repräsentierten Operatoren ist nicht T),
2. die Relationen dieselbe Argumentzahl haben³⁰ und
3. sowohl für das Paar der Anfangsknoten der verknüpften Kanten als auch für das Paar der Endknoten der verknüpften Kanten ein Eintrag in der Matrix existiert.

Durch Punkt 3 wird garantiert, dass die Bedingung bezüglich der Menge der Kantenpaare aus Definition 4.4.1 erfüllt ist. Die Prozedur BILDEKNOTENMATRIX bildet die oben beschriebene Matrix aus den Knoten der Graphen. Die Prozeduren ANFANG und ENDE haben Kanten als Argumente und geben den Anfangsknoten und den Endknoten der Kante zurück. Die mittels

³⁰ Die Prüfung der Stelligkeit der Relationen sollte eigentlich schon mit Punkt 1 überprüft worden sein und sei hier nur der Vollständigkeit halber noch einmal mit aufgeführt.

des folgenden Algorithmus berechnete Verknüpfung wird im Folgenden Produktverknüpfung genannt.

Algorithmus 4.5.1: Gegeben seien zwei CRIL-Graphen $A = (V_A, E_A)$ und $B = (V_B, E_B)$. Berechnet wird eine Produktverknüpfung.

Datenstrukturen:

Verknüpfung: P
 Matrix: knotenMatrix
 CRIL-Knoten: k_A, k_B
 CRIL-Kanten: r_A, r_B
 Paare: k_P, r_P

Prozedur:

```

PRODUKTVERKNÜPFUNG(A, B)
1. knotenMatrix := BILDEKNOTENMATRIX(A, B)
2. for all  $k_A \in V_A$  do
3.     for all  $k_B \in V_B$  do
4.          $k_P := \text{VERKNÜPFE}(k_A, k_B)$ 
5.         KNOTENEINFÜGEN(P,  $k_P$ )
6.         knotenMatrix[ $k_A$ ][ $k_B$ ] :=  $k_P$ 
7.     end
8. end
9. for all  $r_A \in E_A$  do
10.    for all  $r_B \in E_B$  do
11.         $r_P := \text{VERKNÜPFE}(r_A, r_B)$ 
12.        if  $r_P \neq \text{nil}$  do
13.            if (knotenMatrix[ANFANG( $r_A$ )][ANFANG( $r_B$ )]  $\neq$  null  $\wedge$ 
                knotenMatrix[ENDE( $r_A$ )][ENDE( $r_B$ )]  $\neq$  null) do
14.                RELATIONEINFÜGEN(P,  $r_P$ )
15.            end
16.        end
17.    end
18. end
19. return P
  
```

Die gesuchte Verknüpfung, bezüglich der sich die größte Ähnlichkeit der verglichenen Graphen ergibt, ist Teil der Produktverknüpfung. Die gesuchte Verknüpfung ist ferner eine konsistente Verknüpfung. Da die Kantenpaare bereits konsistent sind, müssen also inkonsistente Knotenpaare aus der Produktverknüpfung eliminiert werden. Wenn man nun die inkonsistenten Knotenpaare so eliminiert, dass der Ähnlichkeitsgehalt der Verknüpfung so wenig wie möglich verringert wird, muss das zwangsläufig zu der Verknüpfung führen, die konsistent ist und bezüglich der sich die größte Ähnlichkeit der verglichenen Graphen ergibt. Ein solches Vorgehen entspricht

einem A*-Algorithmus, in dem nicht eine Kostenfunktion minimiert wird, sondern eine Bewertungsfunktion maximiert.

Der A*-Algorithmus garantiert das Finden der besten Lösung, wenn folgende Bedingungen erfüllt sind:

1. Jeder Zustand hat nur endlich viele Folgezustände.
2. Die Kostenfunktion ist immer größer einem bestimmten Wert. Da sich die Kosten eines Folgezustands aus den Kosten des Vorgängers und der Kostenfunktion ergeben folgt daraus, dass die Kosten eines Folgezustands niemals unter denen seines Vorgängers liegen.
3. Die Heuristikfunktion überschätzt für keinen Zustand die Folgekosten bis zum Erreichen eines Zielzustands.

Der Unterschied des im Folgenden präsentierten Algorithmus zum A*-Algorithmus ist, dass nicht die Kostenfunktion minimiert, sondern die Bewertungsfunktion maximiert werden soll. Weiterhin werden in jedem Schritt Knoten- und Kantenpaare entfernt, was zu einer Verkleinerung der Bewertung führt. Will man nun das Finden der besten Lösung garantieren, so bedeutet dies folgendes:

Punkt 1 muss erfüllt sein, was angesichts der endlichen Menge von Operationen zum Bestimmen eines Folgezustands gegeben ist. Die Bewertungsfunktion darf, im Gegensatz zu Punkt 2, niemals eine höhere Bewertung für einen Folgezustand ergeben und die Heuristik darf die Bewertung der Folgezustände niemals unterschätzen. Dadurch wird garantiert, dass ein Zustand niemals eine kleineren (Gesamt-) Bewertung hat, als ein Zielzustand, der über den Zustand erreicht werden kann. Nimmt man nun keine Heuristik und stellt folgende Forderung an die Bewertungsfunktion der Verknüpfungen, so kann das Finden der optimalen Lösung garantiert werden.

Eine Bewertungsfunktion für Verknüpfungen muss mindestens die folgenden Eigenschaften erfüllen:

1. *Das Entfernen eines Knoten- oder Kantenpaares darf die Bewertung nicht erhöhen.*
2. *Das Entfernen von Informationen eines Knoten oder einer Kante (z.B. durch Generalisierung) darf nicht zu einer Erhöhung der Bewertung führen.*

Der Startzustand der Suche ist die durch Algorithmus 4.5.1 berechnete Verknüpfung. Jeder Folgezustand der Suche ist ein Teil des Vorgängers, gewonnen durch Kopieren der Verknüpfung³¹ und Löschen von inkonsistenten

³¹ Verknüpfungen entsprechen in ihrer Struktur Graphen. Kopieren von Graphen kann sehr hohen Berechnungsaufwand haben. In [PoC95] wird als Lösung dieses Problems die Benutzung von virtuellen Graphen vorgeschlagen. Ein virtueller Graph hat für Knoten und Kanten eines Ursprungsgraphen einen Bitvektor, der angibt ob im virtuellen Graphen der jeweilige Knoten oder die jeweilige Kanten enthalten ist oder nicht. Dieses Vorgehen

Knotenpaaren und der Kantenpaare, in denen Kanten enthalten sind, deren Anfangs- oder Endknoten Element des gelöschten Knotenpaares ist. Die so gewonnenen Verknüpfungen werden in einer Liste, sortiert nach ihren Bewertungen (der Ähnlichkeitsgehalt), gespeichert. Die Verknüpfung mit der höchsten Bewertung wird aus der Liste entnommen und auf Konsistenz geprüft. Ist sie konsistent, so ist garantiert, dass sie die gesuchte Verknüpfung ist und der Algorithmus terminiert. Ist sie inkonsistent, wird auf sie die gleichen Prozedur angewendet.

Für den Algorithmus muss bestimmt werden können, ob zwei gegebene Knotenpaare inkonsistent zueinander sind.

Definition 4.5.2: Gegeben seien zwei Knotenpaare $p_1 = (k_1, k_1')$ und $p_2 = (k_2, k_2')$.

$$\text{inkonsistent}(p_1, p_2) \Leftrightarrow p_1 \neq p_2 \wedge [k_1 = k_2 \vee k_1' = k_2']$$

Nach dem Kopieren des Graphen wird ein, mindestens mit einem anderen Knoten, inkonsistentes Knotenpaar gesucht. In der einen Kopie wird eben dieses Knotenpaar gelöscht, in der anderen Kopie alle mit ihm inkonsistenten Knoten (das Knotenpaar wird *durchgesetzt*). Die Prozedur SETZEKNOTENDURCH löscht alle nach Definition 4.5.2 mit dem Knotenpaar inkonsistenten Knotenpaare im Graphen. Die Prozedur LÖSCHE löscht ein Knotenpaar aus der Verknüpfung sowie alle Kantenpaare, in denen ein Paar aus Anfangs- bzw. Endknoten dem gelöschte Knotenpaar entspricht.

Prozedur 4.5.3: SETZEKNOTENDURCH(Verknüpfung G, CRIL-Knoten m)

1. **for all** $n \in K_G$ **do**
2. **if** $\text{inkonsistent}(m, n)$
3. LÖSCHE(G , n)
4. **end**

Um während der Suche zu entscheiden welches Knotenpaar durchgesetzt wird, wird die Prozedur WÄHLEKNOTEN verwendet. Die Prozedur WÄHLEKNOTEN sucht das Knotenpaar der Verknüpfung, das mit den meisten anderen Knotenpaaren inkonsistent ist.

Der folgende Algorithmus berechnet nun die konsistente Verknüpfung zweier Graphen A und B, bezüglich der die Bewertungsfunktion maximal ist. Die Prozedur SORTIEREBZGL sortiert eine Liste bezüglich einer als zweites Argument angegebenen Funktion, wobei die Elemente der Liste Elemente des Definitionsbereichs der Funktion sein müssen. Die Funktion muss die oben bestimmten Eigenschaften erfüllen. Diese Eigenschaften werden von der

ermöglicht es Graphen, (oder Verknüpfungen) sehr effizient zu kopieren, solange die Kopien Teilgraphen des Ursprungsgraphen sind.

Kapitel 5

Repräsentation von Routenwissen

In Kapitel 6 soll gezeigt werden, wie in dem aus der Instruktion stammenden CRIL-Graphen der zu einem bestimmten Zeitpunkt der Navigation relevante Teilgraph gefunden werden kann und wie damit eine Selektion der Perzeption erreicht werden kann. Um den relevanten Teilgraphen zu finden, muss bekannt sein, mit welchen Konzepten räumliches Wissen in CRIL-Graphen repräsentiert wird. Diese Konzepte sollen in diesem Abschnitt vorgestellt werden. Zunächst wird in Abschnitt 5.1 ein allgemeines Modell zur Repräsentation von Routenwissen vorgestellt. In Abschnitt 5.2 wird gezeigt, wie Wissen einer Wegbeschreibung (Routeninstruktion³²) repräsentiert werden kann. Abschnitt 5.3 fasst die Konzepte zusammen und zeigt, wie sie mithilfe von CRIL-Graphen repräsentiert werden können. Die Konzepte, die zur Repräsentation von Wissen einer Routeninstruktion verwendet werden, entsprechen nicht alle den Konzepten, die zur Repräsentation von Wissen aus der Perzeption verwendet werden. Trotzdem müssen I-Graph und P-Graph verglichen und koreferenziert werden können. Dass dies nicht ohne weiteres möglich ist, wird bereits in Abschnitt 5.2 deutlich werden. In Abschnitt 5.4 wird ein Ansatz vorgestellt, wie dieses Problem gelöst werden kann.

5.1 Routen

Meist beinhaltet eine Route mehrere Zwischenstationen. Die Zwischenstationen sind wiederum Orte der Umgebung, an denen es mehrere Möglichkeiten der Weiterreise gibt, an denen also Entscheidungen getroffen werden müssen, in welche Richtung sich der Navigator bewegen muss, um sein

³² Einige Autoren unterscheiden zwischen Routeninstruktion und Routendeskription. Im Allgemeinen enthalten Wegauskünfte aber sowohl instruktive als auch deskriptive Elemente (vgl. z. B. [DaD98]). Ich werde im Folgenden von Routeninstruktionen sprechen, wobei diese ebenfalls deskriptive Elemente enthalten können.

Ziel zu erreichen. Die Route ist also in mehrere Segmente unterteilt, wobei der Endpunkt eines Segments immer der Startpunkt des folgenden Segments ist und an den Start- bzw. Endpunkten taktische Entscheidungen getroffen werden müssen. Start- und Endpunkte entsprechen (offensichtlich) Plätzen von kognitiven Karten, Segmente entsprechen den topologischen Relationen der kognitiven Karten (vgl. Abschnitt 1.1). Von Werner, Krieg-Brückner und Herrmann wird in [WKH00] ein Modell vorgestellt, das für die Repräsentation von Routenwissen verwendet werden kann. Dieses Modell soll als Grundlage für die Repräsentation von Routenwissen in den P-Graphen verwendet werden und wird deshalb im Folgenden vorgestellt.

Eine Route lässt sich nach Werner et al. durch eine Konkatenation von gerichteten **Routensegmenten** repräsentieren. Jedes Routensegment hat einen **Startpunkt** und einen **Endpunkt**. Durch die Angabe eines Start- und Endpunkts wird die Richtung eines Routensegments repräsentiert. Diese Repräsentation der Richtung ist nötig, da Routenwissen immer in Bezug auf die Richtung gelernt und auch mental repräsentiert wird, in der die Route erlernt und repräsentiert wurde. Dieses Wissen über den Verlauf der Route ist nicht unbedingt umkehrbar (vgl. zum Beispiel [WKH00]). Jedes Routensegment enthält Informationen darüber, wie der Navigator dem Routensegment folgen kann. Diese Informationen können zum Beispiel dadurch spezifiziert sein, dass das Routensegment von einer bestimmten Klasse ist wie etwa Bahnlinie oder Straße (zum Beispiel repräsentiert durch eine Sorte). Handelt es sich um eine Bahnlinie, kann dem Routensegment durch eine Bahnfahrt auf der entsprechenden Linie gefolgt werden, handelt es sich um eine Straße, kann ihm durch eine Autofahrt, einen Fußmarsch o. ä. gefolgt werden. Außerdem muss am Startpunkt die Information repräsentiert sein, wie dieses Routensegment eindeutig identifiziert werden kann.

Die Routensegmente sind durch **Plätze** verbunden. An den Plätzen müssen taktische Entscheidungen getroffen werden (vgl. oben), was zu dem von Dehm in [Deh03] eingeführten Begriff des Entscheidungspunktes führt.

Ein **Entscheidungspunkt (EP)** in der Umgebung ist ein Punkt einer Route, an dem der Navigator eine taktische Entscheidung bezüglich des weiteren Routenverlaufs treffen muss oder eine Reorientierung vornehmen muss.

In einer urbanen Umgebung sind also alle Arten von Kreuzungen und Abzweigungen Entscheidungspunkte. Aber auch Konstellationen von Wegen, Straßen oder ähnlichen Routensegmenten, die einer Biegung entsprechen, fallen unter Umständen in die Kategorie der Entscheidungspunkte. An dieser Stelle sei angemerkt, dass eine in einer Routeninstruktion beschriebene Biegung auf jeden Fall ein möglicher EP ist. Ob eine Biegung während der Navigation als EP eingestuft wird, hängt von verschiedenen Faktoren ab. Auf dieses Problem wird im Folgenden noch genauer eingegangen.

Ein Platz ist nach [WKH00] durch eine **Position**, ein **Referenzsystem** und eine **Referenzrichtung** repräsentiert. Ein Platz ist so durch die Lokalisierung

eines Punktes des Platzes (der Position) in Relation zu bestimmten Landmarken und in Bezug auf ein Referenzsystem (mit einer Referenzrichtung) repräsentiert. Zu beachten ist, dass es sich keinesfalls um einen genau definierten Punkt der Umgebung im mathematischen Sinn handelt. Vielmehr gibt es eine Region des Platzes, die relativ zu bestimmten Landmarken in Bezug auf das Referenzsystem bestimmt ist. Der durch die Position repräsentierte Punkt liegt irgendwo innerhalb dieser Region. Wie Position und Referenzsystem in der natürlichen Sprache bestimmt werden, wird in Abschnitt 5.2 erläutert. Positionen sind die Start- und Endpunkte der Routensegmente.

Ein Ort der Umgebung kann durch verschiedene Plätze in verschiedenen Routen repräsentiert sein. Für jede Route, in der der Ort als Platz repräsentiert wird, hat die Repräsentation ein eigenes Referenzsystem. Diese Repräsentationen der Plätze werden so lange als verschieden angesehen, bis ihre Referenzsysteme zu einem integriert werden und folglich auch ihre Richtungen bezüglich diesem einem Referenzsystem dargestellt werden. Eine solche Integration entspricht dem bereits in Abschnitt 1.1 erwähnten Übergang von Routenwissen zu Überblickswissen.

Es sei noch angemerkt, dass es auch andere Möglichkeiten der Repräsentation von Position und Richtung gibt als die eben erläuterte. So deuten Untersuchungen darauf hin, dass bestimmte Tiere Momentaufnahmen der Umgebung an einem Platz speichern. Bei der nächsten Entscheidung an diesem Ort versuchen die Tiere, in eine Position und (Aus-) Richtung zu kommen, so dass das wahrgenommene Bild wieder mit dem Gespeicherten übereinstimmt (vgl. zum Beispiel [WKH00]).

5.2 Routen in Wegbeschreibungen

Das durch eine Routeninstruktion kommunizierte Wissen entspricht Routenwissen. Als solches enthält eine Routeninstruktion die eben eingeführten Repräsentationen von Plätzen und Routensegmenten. Ein I-Graph repräsentiert das räumliche Wissen einer Routeninstruktion, also die Repräsentationen der Plätze und Routensegmente. Um das Wissen adäquat repräsentieren zu können, ist es notwendig zu wissen, in welcher Form das Wissen über Plätze und Routensegmente in Routeninstruktionen enthalten ist. Deshalb soll im Folgenden erläutert werden, wie Plätze und Routensegmente in Routeninstruktionen beschrieben werden.

5.2.1 Plätze

Plätze sind auch in Routeninstruktionen durch eine Position, ein Referenzsystem und eine Referenzrichtung repräsentiert. Das prominenteste Referenzsystem in Routeninstruktionen ist das deiktische Referenzsystem des Navigators. Dabei wird die letzte Bewegungsrichtung als Referenzrichtung verwendet (vgl. z. B. [Kle79] Abschnitt 5.2). Das folgende, aus einer Wegbeschreibung in [Kle79] entnommene Beispiel soll dies verdeutlichen.

(5.1) „Jetzt geh Sie vor, bis Sie auf den Kaufhof stoßen, ... , dann gehen Sie links rein...“

In dem Beispiel wird implizit die letzte Bewegungsrichtung des Navigators für den Aufbau des deiktischen Referenzsystems benutzt und die Identifizierung des nächsten Routensegments („... , dann gehen Sie links rein...“) anhand dieser Richtung vorgenommen. In dem Beispiel ist implizit auch die Angabe einer Position enthalten, nämlich einer Position vor dem Kaufhof. Eine andere Beschreibung aus [Kle79] lautet folgendermaßen:

(5.2) „... vor bis zum Kaufhof, rechts ist der Kaufhof, ja? Un da halten Sie sich rechts...“

In (5.2) ist die Position wieder vor dem Kaufhof lokalisiert. Hierbei ist zu beachten, dass die räumliche Relation keineswegs in der Beschreibung enthalten ist, sondern aus der Tatsache erschlossen wird, dass man auf den Kaufhof zukommt. Die Referenzrichtung wird in Beispiel (5.2) explizit mit angegeben. Sie ist so zu wählen, dass sie der Ausrichtung des Agenten entspricht, wenn der Kaufhof rechts vom Agenten bezüglich des deiktischen Referenzsystems des Agenten liegt. Die Referenzrichtung in Routeninstruktionen ist, wie bereits deutlich wird, immer die Ausrichtung des Navigators. Dementsprechend muss die Referenzrichtung im I-Graphen nicht explizit mit angegeben werden. In Beispiel (5.2) wird aber deutlich, dass es zum Teil nötig ist, dass der Agent sich einer bestimmten Weise ausrichtet, also nicht einfach die letzte Bewegungsrichtung des Agenten als Referenzrichtung verwendet wird. Der Agent wird also (implizit) angewiesen, sich neu auszurichten. Solche Anweisungen sind aber im Aktionsplan des GA enthalten und sollen deshalb an dieser Stelle nicht weiter berücksichtigt werden. Es wird im Folgenden davon ausgegangen, dass der Agent immer entsprechend der Routeninstruktion ausgerichtet ist und damit die Referenzrichtung eindeutig definiert ist.

Im Falle der beiden Beispiele (5.1) und (5.2) muss der durch die Position repräsentierte Punkt irgendwo vor dem Kaufhof lokalisiert sein. Ein sehr prominenter Fall einer Lokalisierung der Position (in Routeninstruktionen) ist noch, dass sich die Position irgendwo *bei* der Einmündung des letzten Routensegments befindet und die Position in der Routeninstruktion nicht explizit lokalisiert wird. Die Aussage „...biege an der nächsten Kreuzung links

ab...“ ist ein Beispiel dafür. Die Lokalisierung der Position ist nicht explizit in der Aussage beschrieben. Einem menschlichen Navigator ist aber klar, dass er, sobald er die Kreuzung erreicht hat, nach einem Routensegment in seiner Umgebung suchen muss, das relativ zu seiner momentanen Position und seiner letzten Bewegungsrichtung links von der Kreuzung abgeht. Die Position ist also bei der Einmündung des Routensegments, das seinen Endpunkt in dem Platz hat. Wenn also keine expliziten Angaben über Position und Ausrichtung in einer Wegbeschreibung gegeben sind, werden Defaultannahmen getroffen. Dies soll auch für den GA und den I-Graphen gelten. Ist in der Routeninstruktion keine explizite Aussage über Position und Richtung gemacht, wird auch keine Information darüber im I-Graphen repräsentiert. Der GA benutzt dann sein Wissen, um zu bestimmen, wo die Position lokalisiert ist. Damit soll die Betrachtung von Plätzen in Routeninstruktion abgeschlossen sein. Es sei festgehalten, dass die Beschreibung von Position, Referenzsystem und Referenzrichtung eines Platzes, zum Teil nur implizit, auch in Routeninstruktionen und damit auch in I-Graphen enthalten sind.

Die durch Routeninstruktionen kommunizierten Plätze sind Entscheidungspunkte. Dabei gibt es verschiedene Arten von Entscheidungspunkten in Routeninstruktionen. Die Unterscheidung der Entscheidungspunkte betrifft die Mechanismen, die nötig sind, um den EP in der Umgebung zu identifizieren. Dehm unterscheidet zwischen den folgenden drei Arten von Entscheidungspunkten:

Der Entscheidungspunkt Landmarke (EPLM) befindet sich in unmittelbarer Nähe einer oder mehrerer Landmarken. Dabei ist die Angabe von räumlichen Relation zu Landmarken eingeschränkt möglich. (Bsp.: „An der Aral-Tankstelle rechts abbiegen.“)

Der Entscheidungspunkt Zählung und Distanzangabe (EPZD) wird durch Abzählen von Landmarken oder Entfernungen identifiziert.³³ (Bsp.: „An der dritten Kreuzung rechts abbiegen.“)

Der Entscheidungspunkt Suche (EPS) kann nur durch kontinuierliche Aufmerksamkeit identifiziert werden. (Bsp.: „Nachdem du einmal um den See herum gegangen bist, geradeaus weiter.“, „Nach der Aral-Tankstelle die nächste Möglichkeit rechts abbiegen.“) Hierbei wird also vom Agenten gefordert, dass er über einen längeren Zeitraum nach dem Entscheidungspunkt sucht.

Die Tatsache, dass es nur bei einem EPS nötig ist, kontinuierlich zwischen zwei Entscheidungspunkten nach Landmarken oder anderen Entscheidungshilfen zu suchen, ist wegen der normalerweise begrenzten

³³ Dehm spricht hier von Wegeknoten. Wegeknoten sind m. E. in diesem Zusammenhang aber auch als Landmarken anzusehen, weshalb ich hier von seiner Terminologie abweiche.

Ressourcen eines künstlichen Agenten wie dem GA äußerst nützlich. Ein menschlicher Agent perzipiert dagegen immer (also auch bei einem EPLM oder EPZD) kontinuierlich, was ihm aufgrund seiner extremen kognitiven Leistungsfähigkeit keinerlei Probleme bereitet.

Es sei noch angemerkt, dass es bei der Identifizierung von Entscheidungspunkten in einer Routeninstruktion zu Verwechslungen zwischen dem EPS und EPLM kommen kann. So kann die Aussage „...nach der Aral-Tankstelle rechts abbiegen...“ sowohl ausdrücken, dass direkt hinter der Tankstelle eine Kreuzung o. ä. liegt (EPLM), aber auch, dass die nächste Möglichkeit rechts abzubiegen, erst in einiger Entfernung hinter der Tankstelle liegt (EPS). An dem Beispiel wird bereits deutlich, dass die in Routeninstruktionen vorkommenden Entscheidungspunkte nicht alle auf der Route in der Umgebung vorkommenden Plätze repräsentieren.

5.2.2 Pfade und Routensegmente

Routensegmente in Routeninstruktionen entsprechen nicht unbedingt den Routensegmenten der Umgebung, in der die beschriebene Route liegt. In Routenbeschreibungen wird zum Beispiel nicht immer explizit deutlich gemacht, von welcher Klasse ein Routensegment ist und es wird nicht unbedingt jeder Platz der Route erwähnt. Vielmehr wird durch die Bewegungsverbren in Routeninstruktionen ein so genannter **Pfad** impliziert, der einem oder mehreren Routensegmenten in der Umgebung entspricht. Im Folgenden soll zunächst kurz beschrieben werden, welche linguistischen Elemente Routeninstruktionen enthalten und wie Bewegungsverbren Pfade charakterisieren. Im Anschluss werden Pfade und Routensegmente miteinander verglichen und gezeigt, dass Routensegmente teile von Pfaden sein können und umgekehrt.

Sätze einer Routeninstruktion beinhalten (unter anderem) Aktionen, Bewegungsdeskriptionen, Positionen und Landmarken (vgl. [TSH03]). Dabei charakterisieren verschiedene Gruppen von Wörtern diese Komponenten. Aktionen werden zum Beispiel durch die Verben der Sätze ausgedrückt. Die Verben werden dabei wiederum in drei verschiedene Gruppen unterteilt bezüglich Positionierung, Bewegung und Orientierungswechsel.

Positionsverbren sind zum Beispiel *sein* oder *stehen*. Sie implizieren, dass sich ein durch den Satz bestimmtes Objekt an einer Position befindet. Verben der Orientierungsänderung, wie zum Beispiel *abbiegen* oder *umdrehen*, implizieren, dass der Träger der Orientierungsänderung, im Fall von Routendeskriptionen der Agent, seine Frontachse neu ausrichtet³⁴.

Bewegungsverbren wie *gehen*, *laufen* oder *betreten* bestimmen einen so genannten Bewegungsträger und einen so genannten **Pfad** (vgl. [ETH00]). Dabei impliziert das Verb, dass sich der Bewegungsträger entlang des Pfades bewegt. Die Verben können noch bezüglich der implizierten Bewegungsart

³⁴ Für eine geometrische Charakterisierung von Orientierungsänderungen vgl. z.B. [STE00].

unterschieden werden. Allerdings werden in Routeninstruktionen häufig Verben verwendet, die (relativ) neutral bezüglich der Bewegungsart sind. Für den Geometrischen Agenten ist eine Unterscheidung der Bewegungsart schon von daher nicht von Interesse, da es ihm unmöglich ist, sich auf verschiedenen Weisen zu bewegen. Die Repräsentation der Aktionen ist Teil des Aktionsplans. Sie ist für die vorliegende Arbeit deshalb nicht von Interesse und soll nicht weiter betrachtet werden.³⁵

Von Interesse für die vorliegende Arbeit ist aber eine geometrische Spezifikation von Pfaden, da sie ein Teil der räumlichen Komponente der Routeninstruktionen sind. So eine Spezifikation wird in [ETH00] vorgestellt und soll hier kurz zusammengefasst werden. Pfade repräsentieren den räumlichen Anteil von Kombinationen aus Verben und Richtungspräpositionen, wie sie in Routeninstruktionen häufig verwendet werden. Beispiel hierfür sind „gehe in das Haus“, „biege nach links ab“. Diese Kombinationen implizieren eine begrenzte, lineare, gerichtete Struktur, entlang derer sich der Bewegungsträger bewegt. Diese Struktur wird von Eschenbach et al. in [ETH00] Pfad genannt. Dass die Geometrie des Pfades auch vom Verb allein abhängen kann, lässt sich aus der Tatsache ableiten, dass auch Sätze, die keine Richtungspräposition enthalten wie z. B. „Thomas wirft den Ball“ dem Bewegungsträger (im Beispiel der Ball) einen Pfad zuordnen. Das Verb verbindet dabei den Bewegungsträger mit dem Pfad. Im Normalfall aber charakterisiert die Richtungspräposition die räumliche Struktur des Pfades. Präpositionen wie *in*, *an* oder *zu* implizieren, dass der Pfad einen Endpunkt hat, der in der Region liegt, die durch die Präposition bestimmt ist. Der Startpunkt des Pfades liegt aber grade nicht in dieser Region. Präpositionen wie *aus* oder *von* implizieren, dass der Pfad einen Startpunkt hat, der in der durch die Präposition bestimmten Region liegt, während der Endpunkt nicht in der Region liegt. Die Eigenschaft, einen Start- und einen Endpunkt zu haben, drückt gleichzeitig die Richtung des Pfades aus. Die Präposition *durch* impliziert, dass der Pfad einen Innenpunkt hat, also einen Punkt, der auf dem Pfad liegt, aber weder Start- noch Endpunkt ist, welcher in einer Region liegt. Die Region ist durch eine zusätzliche Beschreibung, wie etwa die Präposition *zwischen* oder das Objekt *Tür*, bestimmt. Für eine formelle Spezifikation der Geometrie von Pfaden sei wieder auf [ETH00] verwiesen. Die eben eingeführten Punkte eines Pfades werden, wie die einen Platz charakterisierenden Punkte, in CRIL-Graphen durch Positionen repräsentiert. Diese Positionen dienen so der Lokalisierung der Pfade.

Die Bewegungsverben werden im I-Graphen durch Pfade repräsentiert. Ein Pfad muss also einen Start- und einen Endpunkt haben und kann einen Innenpunkt haben. Die strukturelle Ähnlichkeit zwischen Pfaden und Routensegmenten ist auffällig. Beides sind lineare gerichtete Entitäten, die einen Start und einen Endpunkt haben. Allerdings sind die Segmentierungen von Pfaden und Routensegmenten nicht unbedingt gleich, wie im Folgenden gezeigt werden wird.

³⁵ Für eine eingehendere Analyse von Bewegungsverben vgl. [ETH00].

Feststellung 5.2.1: *Pfade repräsentieren Routensegmente, einen Teil von Routensegmenten oder mehrere Routensegmente in Routeninstruktionen.*

Dass Feststellung 5.2.1 gilt, sollen folgende Beispiele verdeutlichen:

- (5.3) „Gehe von hier (geradeaus) bis zur nächsten Kreuzung.“
- (5.4) „Wenn du aus der Mensa kommst, bieg nach links ab, ...“
- (5.5) „Gehe die Straße entlang, bis du zur Aral-Tankstelle kommst.“

Die Aussage aus Beispiel (5.3) repräsentiert genau ein Routensegment³⁶. Dabei sind die beschriebenen Kreuzungen die Plätze bzw. Entscheidungspunkte und die Straße zwischen den Kreuzungen ist das Routensegment.

In der Aussage von Beispiel (5.4) sind zwei Pfade enthalten, nämlich ein Pfad w_1 , der seinen Startpunkt in oder bei der Mensa hat und einen Pfad w_2 , der seinen Endpunkt in der Linksregion bezüglich w_1 hat.³⁷ Der Endpunkt von w_2 stellt aber nicht den Endpunkt eines Routensegments dar, sondern vielmehr einen beliebigen Punkt des Routensegments. Er repräsentiert keinen Entscheidungspunkt, sondern dient nur der Lokalisierung des Pfades w_2 und somit des Routensegments, das w_2 enthält.

In Beispiel (5.5) besteht der repräsentierte Pfad unter Umständen aus mehreren Routensegmenten. Die Aussage impliziert, dass der Navigator der Straße auch dann folgen soll, wenn sie Kreuzungen oder andere Entscheidungspunkte enthält. Der Pfad steht also für mehrere Routensegmente.³⁸ Die Beispiele machen deutlich, dass Feststellung 5.2.1 gilt und weshalb sie gilt.

5.3 Räumliches Inventar der CRIL

Als Zusammenfassung der vorangegangenen beiden Abschnitte wird im Folgenden dargestellt, welches Inventar an Sorten und Relationen CRIL-Graphen benötigen, um die Routen zu repräsentieren. Dabei werden zunächst die benötigten Sorten der CRIL-Knoten beschrieben und im Anschluss die Relationen, die durch Kanten zwischen Knoten der entsprechenden Sorten repräsentiert sind.

³⁶ Es repräsentiert genau genommen nur dann ein Routensegment, wenn es bis zur nächsten Kreuzung keine weiteren Möglichkeiten gibt, die Straße zu verlassen, es also keine Möglichkeit gibt, eine taktische Entscheidung bezüglich der Weiterreise zu treffen. Es sei an dieser Stelle angenommen, dass dies der Fall ist. Es wird aber schon deutlich, dass eine Zuordnung von Pfaden auf Routensegmente Schwierigkeiten bereitet.

³⁷ Im Folgenden stehen Variablen, die mit w beginnen, für Pfade.

³⁸ Auch hier besteht natürlich wieder die Möglichkeit, dass nur ein Routensegment durch die Aussage repräsentiert wird. Außerdem wird hier deutlich, dass nicht nur ein mentales Objekt eines Pfades in einer Routeninstruktion enthalten sein kann, sondern auch Objekte der Umgebung. In diesem Fall ist dieses Objekt (der Umgebung) die Straße.

5.3.1 Sorten

Für die Repräsentation der räumlichen Information, die in Routeninstruktionen enthalten ist, müssen Pfade repräsentiert werden. Das Konzept des Pfads wurde bereits eingeführt. Es repräsentiert eine (lineare) gerichtete räumliche Struktur, entlang der sich der Navigator bewegen kann und soll. Es wird folglich eine Sorte Pfad³⁹ benötigt. Start- und Endpunkte der Pfade werden im CRIL-Graphen durch Knoten der Sorte Position repräsentiert.

Das Gegenstück des Pfads in der Umgebung der Route ist ein Objekt der Umgebung, das die Geometrie des Pfades repräsentiert. Im Falle einer vom Menschen gegebenen Wegbeschreibung können Gegenstücke verschiedene Objekte der Umgebung sein, die eine Bewegung des Agenten an ihnen entlang, durch sie hindurch, über sie hinweg oder ähnliches erlauben. Solche Objekte können zum Beispiel eine Tür, ein Trampelpfad oder ein Zaun, eine Autobahn oder eine Bahnverbindung sein. Generell lässt sich sagen, dass es Landmarken sind, denen zu folgen es möglich ist. Solche Strukturen der Umgebung des Agenten sollen durch die Sorte Kurslandmarke repräsentiert werden. In der Umgebung des Agenten sind dies zur Zeit allerdings nur Wege, die durch Knoten der Sorte Weg repräsentiert werden, und Abzweigungen und Kreuzungen von Wegen, die durch Knoten der Sorte Wegkonfiguration repräsentiert werden. Dabei werden die Sorten Weg und Wegkonfiguration in der Taxonomie der Eigenschaften von der Sorte Kurslandmarke subsumiert. Kurslandmarken sollen Routensegmenten in der Umgebung des Agenten entsprechen. Routensegmente brauchen einen Start- und einen Endpunkt. Da Kurslandmarken der Umgebung aber ungerichtet sind, sollen sie zwei Endpunkte haben. Die Endpunkte sind wie die Start- und Endpunkte eines Pfades durch Positionen repräsentiert. Positionen repräsentieren also Punkte der Umgebung des Agenten, sowohl die Positionen des I-Graphen als auch die Positionen des P-Graphen. Da die Routensegmente nach [WKH00] gerichtet sein sollen, muss es aber zusätzlich eine Repräsentation der Richtung geben. Diese Richtung wird durch Pfade repräsentiert, die wiederum die Kurslandmarken charakterisieren. Wie diese Pfade für eine Kurslandmarke bestimmt werden, wird in Abschnitt 5.4 beschrieben.

Wegkonfigurationen stellen Entscheidungspunkte dar. Um einen Entscheidungspunkt zu repräsentieren, wird die gleichnamige Sorte eingeführt. Die Sorte Entscheidungspunkt wird von der Sorte Kurslandmarke subsumiert und benötigt im Gegensatz zu den Kurslandmarken keine Endpunkte. Für die Objektlokalisierung werden ferner Knoten der Sorte Region benötigt. Was Knoten der Sorten Region repräsentieren, sollte bereits deutlich geworden sein. Eine genauere Einordnung in den Kontext der CRIL-Graphen erfolgt in den folgenden Abschnitten.

Damit sind alle grundlegenden, zum Verständnis der nachfolgenden Abschnitte benötigten Sorten eingeführt. Auf die Taxonomie der Sorten möchte ich an

³⁹ Sorten, Relationen und Prädikate werden im Folgenden durch Lucida Font dargestellt.

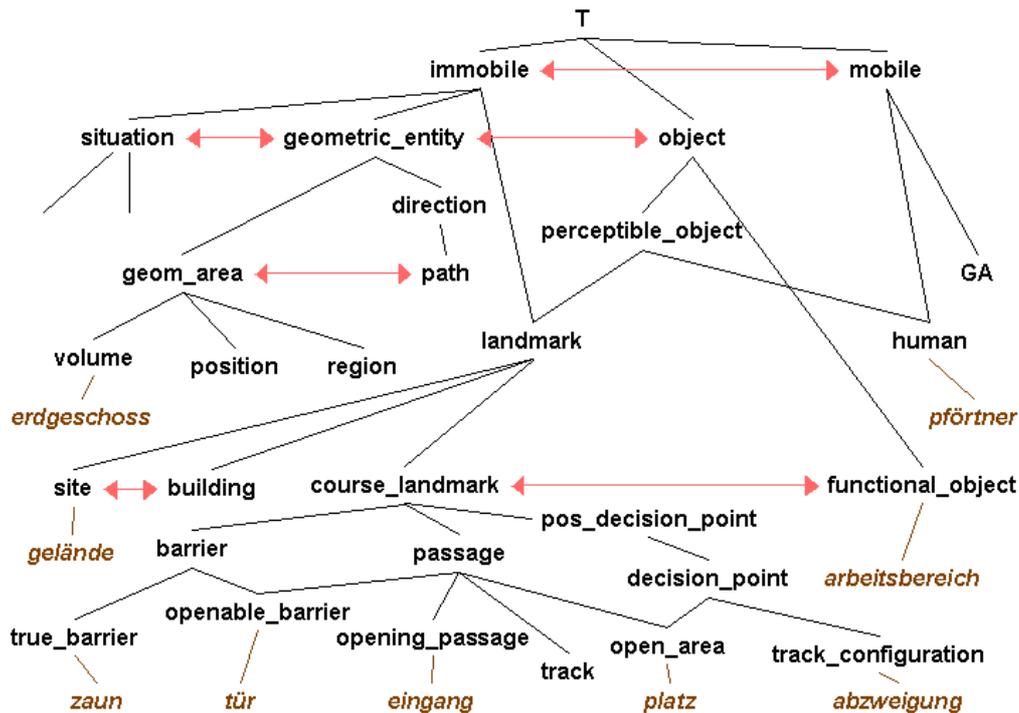


Abbildung 5.1: Ein Beispiel einer Taxonomie von Sorten (schwarz) und Eigenschaften (braun). Die Subsumptionsrelation ist durch schwarze Kanten dargestellt, die Exklusionsrelation durch rote Doppelpfeile. Höher liegende Konzepte subsumieren tiefer liegende Konzepte.

dieser Stelle nicht näher eingehen und sie stattdessen nur informell durch eine Grafik der Hierarchie einführen. Die hier noch nicht eingeführten, aber in der Grafik erscheinenden Sorten, meist Supersorten der hier vorgestellten Sorten, werden hier nicht weiter beschrieben. Ihre Bedeutung ist aber leicht aus der Grafik erschießbar.

5.3.2 Relationen

Wie im vorigen Abschnitt bereits dargestellt wurde, hängen Pfade sowie Kurslandmarken mit Positionen in enger Weise zusammen. Positionen repräsentieren Punkte der Umgebung. Positionen können entweder Start- oder Endpunkte von Pfaden oder Kurslandmarken sein. Dementsprechend muss es Kanten geben, die Relationen $\text{Startpunkt}(w,p)$ und $\text{Endpunkt}(w,p)$ repräsentieren. Dabei ist die Variable w ein Knoten der Sorte Pfad oder Kurslandmarke, und die Variable p ein Knoten der Sorte Position. Durch die Relationen Start- und Endpunkt sind die Richtungen der Pfade bestimmt. Kurslandmarken repräsentieren Objekte der Umgebung, die zur Bewegung des Navigators an ihnen entlang, durch sie hindurch oder über sie hinweg benutzt

werden können. Dabei enden solche Passagen immer in Wegkonfigurationen. Der Endpunkt einer Kurslandmarke ist durch eine Position bestimmt. Die Position wiederum repräsentiert einen Punkt der Wegkonfiguration. Das Gleiche gilt für Position eines I-Graphen die als Entscheidungspunkt identifiziert wurden⁴⁰. Um zu repräsentieren, dass eine Position einen Punkt eines Entscheidungspunkts bzw. einer Wegkonfiguration repräsentiert, wird die Relation $\text{enthält}(ep, p)$ eingeführt, die zwischen einem Entscheidungspunkt ep und einer Position p gelten kann.

Wie bereits erläutert, können Positionen auch der Lokalisierung von Pfaden dienen. Für die Lokalisierung der Positionen und damit auch der Pfade werden Knoten der Sorte `Region` benutzt und es wird eine Relation `lokalisiert(p,r)` eingeführt. Durch die Relation wird ausgedrückt, dass der durch die Position p repräsentierte Punkt in der Region r liegt. Es können grundsätzlich auch andere sichtbare Objekte in einer Region lokalisiert sein. Analog zur Lokalisierung von Position lässt sich zum Beispiel eine beliebige Landmarke, wie ein Haus oder Ähnliches lokalisieren. Generell drückt die Relation `lokalisiert` aus, dass das erste Argument in der Region lokalisiert ist. Um den Verlauf von Pfaden zu repräsentieren, muss es ferner möglich sein, explizit zu repräsentieren, dass eine Position nicht in einer Region lokalisiert ist (vgl. Abschnitt 5.2.2). Zu diesem Zweck wird die Relation `nlokalisiert` verwendet. In einem logikbasierten Repräsentationsformat ließe sich dies auch durch die Verwendung der Negation erreichen. Da die Kanten des CRIL-Graphen aber auf ein bestimmtes, die Relation repräsentierendes Konzept abgebildet werden, ist dies in CRIL-Graphen nicht möglich. Sollte der CRIL-Graph auf eine logikbasierte Repräsentation abgebildet werden, kann die Relation natürlich auf das Prädikat $\neg\text{lokalisiert}(p,r)$ abgebildet werden.

Regionen repräsentieren räumliche Strukturen, die zur Lokalisierung von Objekten oder Punkten der Umgebung verwendet werden. In der natürlichen Sprache werden Objekte durch topologische Relationen und Anordnungsrelationen wie z. B. *vor* oder *in* lokalisiert. Der GA verwendet für die Objektlokalisierung nur Konzepte, die auch in der natürlichen Sprache verwendet werden. Das bedeutet, dass sowohl im I-Graphen als auch im P-Graphen natürlichsprachliche Konzepte für die Repräsentation der Objektlokalisierung verwendet werden. Um diese Konzepte zu verstehen, wird im Folgenden kurz erläutert, wie Objektlokalisierung in der natürlichen Sprache funktioniert. Dabei wird nur die intrinsische und die deiktische Lokalisierung betrachtet. Für eine eingehende Einführung in dieses Thema sei zum Beispiel auf [HeS98] verwiesen.

Objekte werden in der natürlichen Sprache häufig in Bezug auf ein Referenzobjekt⁴¹, eine Origo bzw. ein Referenzsystem und unter Verwendung einer räumlichen Relation lokalisiert. Bei dieser relativen Objektlokalisierung ist das Referenzobjekt das Objekt, bezüglich dessen das lokalisierte Objekt

⁴⁰ Hier sei angemerkt, dass nicht jede Position einem EP entspricht. Auf diesen Punkt wird im Folgenden noch näher eingegangen.

⁴¹ ‚Referenzobjekt‘ hier im Sinne der DI-Dichotomie (vgl. z.B. [HeS98]).

(Lokalisationsobjekt) lokalisiert wird. Folgendes Beispiel soll dies verdeutlichen. Ein Sprecher A teilt einem Empfänger B mit: „Der Ball liegt vor dem Auto.“ In diesem Beispiel ist das Auto das Referenzobjekt und der Ball das Lokalisationsobjekt. Was in diesem Beispiel die Origo ist, ist nicht eindeutig zu bestimmen. So könnte zum Beispiel der Empfänger B die Origo sein, was dem Ausdruck „Von dir aus ist der Ball vor dem Auto.“ (deiktische Lokalisierung) entspricht, oder das Auto ist die Origo, was bedeuten würde, dass der Ball vor dem Auto bezüglich der Fahrtrichtung des Autos liegt (intrinsische Lokalisierung). Bei der intrinsischen Lokalisierung werden die Richtungen, in denen das Lokalisationsobjekt bezüglich des Referenzobjekts liegt, aus bestimmten Eigenschaften des Objekts wie zum Beispiel der Form oder der Funktionalität abgeleitet. Das Beispiel macht deutlich, dass bei natürlichsprachlichen Ausdrücken nicht immer eindeutig bestimmbar ist, welche Origo gemeint ist. Wie und ob solche Mehrdeutigkeiten während der Analyse der Wegbeschreibung aufgelöst werden, soll hier nicht beantwortet werden. Es besteht Grund zur Annahme, dass der Empfänger solcher Aussagen mental eine Region bezüglich des Referenzobjekts bestimmt, in der das Lokalisationsobjekt liegen soll, damit er die Aussage als zutreffend einstuft (vgl. [HeS98] und [Sch99]). Diese Regionen werden in CRIL-Graphen durch Knoten der Sorte Region repräsentiert. Um dem Geometrischen Agenten zu ermöglichen,

- a) die Regionen auch in seiner Umgebung zu bestimmen und damit auch die Objekte, die im I-Graphen als in der Region liegend repräsentiert sind, in der Umgebung zu lokalisieren und
- b) die Lokalisierungen von Objekten der Umgebung, die der Agent während der Navigation vorgenommen hat, adäquat in P-Graphen darzustellen

soll im Folgenden gezeigt werden, wie mithilfe der eben eingeführten Konzepte und Kanten, die räumliche Relationen repräsentieren, Regionen charakterisiert werden können.

Für die Charakterisierung der Regionen, die durch Anordnungsrelationen der natürlichen Sprache bestimmt werden, werden dreistellige Relationen benötigt. Zum Beispiel kann der Ausdruck „Der Ball liegt (von dir aus) vor dem Auto“ durch $\text{vor}(„\text{Auto}“, a_g, r)$ repräsentiert werden, wobei die Variable a_g der Empfänger des Ausdrucks sei und die Variable r eine Region. Um dann zu repräsentieren, dass der Ball in dieser Region liegt, kann die bereits oben eingeführte Relation $\text{lokalisiert}(„\text{Ball}“, r)$ verwendet werden. Also muss für jede Anordnungsrelation der natürlichen Sprache eine dreistellige Relation eingeführt werden. Das erste Argument der Relation sei das Referenzobjekt, das zweite Argument die Origo und das dritte Argument die Region, die durch die Relation bestimmt wird. Weiterhin muss die sprachliche Relation *zwischen* repräsentiert werden können. Als Beispiel soll der Ausdruck „Der Weg liegt zwischen den Häusern B und C“ dienen. In dem Beispiel gibt es zwei Referenzobjekte, die Häuser, und ein Lokalisationsobjekt, den Weg. Die

Verwendung einer Origo ist nicht nötig. Folglich kann *zwischen* durch eine dreistellige Relation repräsentiert werden, wobei die ersten zwei Argumente die Referenzobjekte sind und das dritte die Region ist.⁴²

Die Repräsentation von topologischen Relationen wie zum Beispiel *in* oder *an* ist besonders einfach, da nur ein Referenzobjekt nötig ist. So wird für jede topologische Relation eine zweistellige Relation eingeführt, in der das erste Argument das Referenzobjekt und das zweite Argument die Region ist.

5.4 Koreferenz von Wegen der Instruktion und der Perzeption

Während der Navigation des Geometrischen Agenten müssen die P-Graphen mit dem bestehenden Wissensbestand verglichen und koreferente Elemente der Repräsentationen gefunden werden. In den I-Graphen werden Pfade für die Charakterisierung der Kurslandmarken, also der Routensegmenten verwendet. In den P-Graphen sollen Kurslandmarken die wahrgenommenen Objekte repräsentieren, denen der Agent folgen kann. Für Pfade und Kurslandmarken ist eine direkte Koreferenz nicht möglich. Zum einen da sie generell verschiedene Konzepte repräsentieren und zum anderen aufgrund Feststellung 5.2.1. Für die Integration des P-Graphen in den bestehenden Wissensbestand, also den I-Graphen, bedeutet dies, dass Pfade des I-Graphen und Kurslandmarken des P-Graphen einer besonderen Behandlung bedürfen. Um zu erreichen, dass eine direkte Koreferenz zwischen den Knoten des I-Graphen und den Knoten des P-Graphen möglich ist, sollten die Graphen die gleiche Struktur aufweisen. Wie erreicht werden kann, dass die Struktur der Graphen bezüglich der Kurslandmarken bzw. Pfade gleich ist, ohne die durch die Graphen repräsentierten Informationen zu verändern, soll im Folgenden gezeigt werden.

Die Pfade der I-Graphen charakterisieren wie dargestellt Kurslandmarken. Eine Kurslandmarke soll immer in einem Platz, also einem Entscheidungspunkt, enden. Wenn also bekannt ist, welche Positionen des I-Graphen der Charakterisierung von Entscheidungspunkten dienen, lässt sich in den I-Graphen zwischen je zwei Entscheidungspunkten eine Kurslandmarke einfügen, deren Endpunkte die Entscheidungspunkte sind. Alle Positionen, die Endpunkte oder Innenpunkte der Pfade zwischen den Entscheidungspunkten sind, werden dann als Innenpunkte der Kurslandmarke repräsentiert. Umgekehrt lässt sich für eine Kurslandmarke in einem P-Graphen eine beliebige Anzahl von Pfaden finden, die wiederum die Kurslandmarke charakterisieren. Durch die Verwendung von Pfaden in den P-Graphen lässt

⁴² Genau genommen handelt es sich nicht um zwei Referenzobjekte, sondern um ein komplexes Referenzobjekt, das ein komplexes Diskursobjekt ist. Komplexe Diskursobjekte sind interne Stellvertreter für Objekte der Welt, also in dem angegebenen Beispiel die Häuser. Für eine eingehendere Untersuchung der Präposition *zwischen* sei auf [Hab89] verwiesen.

sich auch die Richtung des durch die Kurslandmarke repräsentierten Routensegments darstellen. Die Frage ist dann, wie viele Pfade eine Kurslandmarke eines P-Graphen charakterisieren sollen. Die Repräsentation einer Kurslandmarke wird während der Navigationsphase des Geometrischen Agenten generiert. Bevor sie generiert wird, wird der I-Graph analysiert und es wird bestimmt, welcher Teil des I-Graphen grade relevant ist (vgl. Abschnitte 6.1 und 6.2). In dem relevanten Teil des I-Graphen ist die relevante Kurslandmarke enthalten. Jede Position der Kurslandmarke des I-Graphen soll einen Punkt einer Kurslandmarke der Umgebung repräsentieren. Deshalb muss jede Repräsentation einer Kurslandmarke der Umgebung im P-Graphen genauso viele Positionen haben wie die Kurslandmarke des I-Graphen. Also wird eine Kurslandmarke der Umgebung durch einen Knoten des P-Graphen repräsentiert, der für jeden End- oder Innenpunkt der Kurslandmarke des relevanten Teilgraphen ebenfalls einen End- oder Innenpunkt hat. Die Struktur der Pfade kann ebenfalls als gleich angesehen werden. So wird also für jeden Pfad des relevanten Teilgraphen ein entsprechender Pfad in den P-Graphen eingefügt, der als Start-, End- und Innenpunkte die entsprechenden Positionen der Kurslandmarke des P-Graphen hat. So lässt sich jede Kurslandmarke eines P-Graphen durch Pfade charakterisieren und die Struktur der Kurslandmarken und Pfade sind in I-Graph und P-Graph gleich. Damit lässt sich ein Vergleich, wie er in Kapitel 4 beschrieben wurde, ohne weiteres durchführen. Dies soll folgendermaßen festgehalten werden, wobei kein Wert auf formelle Korrektheit gelegt wurde:

(5.4.1) *Zwei Positionen p_1 und p_2 , die Entscheidungspunkte charakterisieren und die durch eine unbestimmte Anzahl von Pfaden w_1, \dots, w_n verbunden sind, deren Start- und Endpunkte keine Entscheidungspunkte charakterisieren, es sei denn es handelt sich um p_1 und p_2 , müssen durch eine Kurslandmarke verbunden sein. Die Start-, End- und Innenpunkte der Pfade w_1, \dots, w_n die nicht p_1 und p_2 sind, sind die Innenpunkte der Kurslandmarke.*

Umgekehrt müssen zwei Positionen p_1 und p_2 , die durch eine Kurslandmarke verbunden sind, durch eine unbestimmte Zahl von Pfaden w_1, \dots, w_n verbunden sein, deren Start- und Endpunkte keine Entscheidungspunkte charakterisieren, es sei denn es handelt sich um p_1 und p_2 . Die Start-, End- und Innenpunkte der Pfade w_1, \dots, w_n die nicht p_1 und p_2 sind, sind die Innenpunkte der Kurslandmarke.

Um die Kurslandmarken in die I-Graphen einzufügen, muss aber noch bekannt sein, welche der Positionen Entscheidungspunkte charakterisieren. Dies ist von vornherein nicht unbedingt klar. Zwar werden Entscheidungspunkte zum Teil explizit in Routeninstruktionen erwähnt, aber eben nicht immer. Es stellt sich also die Frage, welche der Start- und Endpunkte von Pfaden eine Position eines Entscheidungspunkts repräsentieren und wie diese im I-Graphen identifiziert werden können. Eine hinreichende Analyse von Routeninstruktionen, die eine eindeutige Identifizierung der Entscheidungspunkte ermöglicht, kann nicht

Gegenstand dieser Arbeit sein. Sie bedarf einer eingehenden linguistischen Analyse. Es soll an dieser Stelle aber trotzdem ein Vorschlag gemacht werden, wie die Suche nach Entscheidungspunkten in einem I-Graphen aussehen könnte.

Der einfachste Fall wird durch Beispiel (5.3) dargestellt. In der Aussage ist von einer Kreuzung und somit von einem Entscheidungspunkt die Rede. Selbstverständlich wird die Information, dass der Endpunkt des Pfades eine Kreuzung ist, in der Analyse der Wegbeschreibung berücksichtigt und im I-Graphen entsprechend repräsentiert. Die Identifizierung des Entscheidungspunktes ist somit trivial.

Schwieriger ist die Identifizierung hingegen bei Beispiel (5.4), in dem kein direkter Hinweis gegeben wird, dass es sich um einen EP handelt. Hier wird aber der zweite Pfad w_2 durch die Anordnungsrelation *links* bezüglich des ersten Pfades w_1 lokalisiert. Für den menschlichen Leser ist dadurch klar, dass hier eine Reorientierung vorgenommen werden muss, also ein EP vorliegt. Der Endpunkt von w_2 hingegen dient nur der Lokalisierung des folgenden Routensegments. Er ist also nicht als ein Entscheidungspunkt zu identifizieren. Die Wegbeschreibung könnte folgendermaßen weitergehen:

(5.6) „... zwischen Haus B und C durch.“

Diese Aussage wird im I-Graphen durch einen Pfad w_3 repräsentiert, dessen Start- und Endpunkt nicht in der Region zwischen Haus B und C liegen, der aber einen Innenpunkt hat, der in der Region liegt. Der intuitive Ansatz wäre nun, den Endpunkt von w_2 gleich dem Startpunkt von w_3 zu setzen. Dies ist m. E. nicht richtig. Der Endpunkt von w_2 dient, wie bereits erläutert, der Lokalisierung von w_2 bzw. der Kurslandmarke in der Umgebung, die durch w_2 repräsentiert wird. Der Startpunkt von w_3 repräsentiert einen möglichen EP. Die Beschreibung, zwischen Haus B und C durch zu gehen, impliziert nämlich nicht zwangsläufig einen EP. Es wäre auch möglich, dass sie dem Navigator nur zeigen soll, dass der eingeschlagene Weg der richtige ist. Es ist bei einer solchen Beschreibung aber davon auszugehen, dass vor der Region zwischen Haus B und C ein EP liegen könnte und die Beschreibung vom Instrukteur so gewählt wurde, damit der Navigator an dem EP den richtigen Weg wählt.

Nun ließe sich argumentieren, dass der Startpunkt von w_3 einem Punkt des Pfades entspricht, der links von w_1 abgeht, und somit die Gleichsetzung des Startpunkts von w_3 mit dem Endpunkt von w_2 korrekt ist. Dies ist sicherlich richtig, wirft aber Probleme beim Erkennen von w_2 auf. Denn wenn der Endpunkt von w_2 als ein möglicher EP klassifiziert wurde, so wird der Agent (zumindest mit erhöhter Priorität) auch nach einem EP suchen. Es ist aber nicht sicher, dass es einen solchen EP überhaupt gibt und wenn, dass dieser von der Position des Agenten aus sichtbar ist. Um nun eine korrekte Darstellung zu gewährleisten, muss zwischen den beiden Pfaden ein zusätzliches Verbindungsstück eingefügt werden. Diese Verbindung sei wiederum ein Pfad mit entsprechendem Start- und Endpunkt und der Kennzeichnung, dass es sich um eine Fortsetzung des vorherigen Pfades handelt. Dann ist der Startpunkt

von w_3 ein möglicher EP. Das ist korrekt, da m. E. die räumliche Beschreibung von w_3 darauf hin deutet, dass es davor einen EP gibt.

Nun lassen sich Entscheidungspunkte und mögliche Entscheidungspunkte folgendermaßen charakterisieren. Damit lässt sich dann die Suche nach Entscheidungspunkten in einem I-Graphen relativ einfach durchführen.

Eine Position ist dann ein EP, wenn

1. sie explizit als ein solcher gekennzeichnet ist oder
2. sie der Startpunkt eines Pfades ist und es ein begehbares Objekt in der Repräsentation gibt, das quer zu ihm liegt, das also explizit als rechtsabgehend, linksabgehend oder quer gekennzeichnet ist.

Ein Startpunkt eines Pfades ist dann ein möglicher EP, wenn

1. es eine Position des Pfades gibt, die nicht der Startpunkt ist und die relativ zu einem beliebigen Objekt lokalisiert wurde oder
2. der Pfad selber die Beschreibung eines neuen begehbaren Objekts, wie z. B. einer Straße, ist.

Eine Position wird also immer dann als EP, sei es ein möglicher oder ein sicherer, identifiziert und markiert, wenn eine räumliche Beschreibung erfolgt oder ein neues begehbares Objekt eingefügt wird. Dies entspricht m. E. dem menschlichen Verständnis, da ein Mensch auch davon ausgehen würde, dass der Instrukteur eine solche Beschreibung äußert, um eine Entscheidungshilfe an einem EP zu geben. Daraus folgt die Möglichkeit, dass es vor dem folgenden Teil der Route einen EP geben kann. Es sei im Folgenden davon ausgegangen, dass die Identifizierung von möglichen oder sicheren Entscheidungspunkten während der Instruktionsphase möglich ist.

Nun wird zwischen je zwei Entscheidungspunkten des I-Graphen eine Kurslandmarke erzeugt, die den bzgl. der Route ersten EP als Startpunkt hat und den bzgl. der Route zweiten EP als Endpunkt hat. Alle anderen Positionen, die zwischen diesen EP liegen, sind Innenpunkte der jeweiligen Kurslandmarke.

Zu beachten ist noch, dass während der Navigation immer auch überprüft werden muss, ob die Innenpunkte der Kurslandmarke aus dem I-Graphen bis zum Erreichen des nächsten korrekt koreferenzierten Entscheidungspunktes ebenfalls mit einem Punkt der Kurslandmarke aus dem P-Graph koreferenziert worden sind. Ist dies nicht der Fall, so lässt das einen der beiden folgenden Schlüsse zu:

1. Der Instrukteur hat sich bezüglich des Innenpunkts geirrt. Der Agent befindet sich aber noch auf dem richtigen Weg.
2. Der Agent hat sich verlaufen. Er muss überlegen, wo er sich vertan haben könnte.

Wie in diesem Fall weiter zu verfahren ist, sei hier nur kurz skizziert, wie das weitere Vorgehen aussehen könnte. Geht man, aufgrund welcher Tatsachen auch immer, davon aus, dass Punkt 1 eingetreten ist, so bietet es sich an, zunächst zu versuchen, „normal“ weiter zu navigieren. Im Normalfall sollte es nicht lange dauern, bis klar wird, ob Punkt 1 oder Punkt 2 eingetroffen ist. Ist Punkt 1 eingetroffen, so sollte es bei den nächsten EP wieder zu einer guten Übereinstimmungen zwischen dem P-Graphen und dem I-Graphen kommen und die Navigation kann normal fortgesetzt werden. Kommt es nicht zu den erwähnten hohen Übereinstimmung, so kann davon ausgegangen werden, dass Punkt 2 eingetroffen ist. In diesem Fall sollte m. E. zu dem letzten EP zurückgekehrt werden und dort nach einem hier nicht beschriebenen Verfahren das nächstbeste Routensegment ausgewählt werden. Selbstverständlich ist ein ähnliches Verfahren anzuwenden, wenn nach einer bestimmten Zeit oder einer bestimmten Anzahl von EP in der Umgebung nicht der bezüglich des I-Graphen korrekte Entscheidungspunkten identifiziert werden konnte. Wann dies der Fall sein soll, hängt mit der hier nicht näher spezifizierten Geduld des Agenten zusammen.

Kapitel 6

Koreferenz in der Navigationsphase

In diesem Kapitel soll ein Vorschlag dargestellt werden, wie mithilfe der in Kapitel 4 eingeführten Ähnlichkeitsmaße die vermuteten Koreferenzen zwischen Knoten eines I-Graphen und Knoten von P-Graphen berechnet werden können. Dafür wird in Abschnitt 6.1 zunächst der grundlegende Ablauf der Perzeption, der Generierung der P-Graphen und der anschließenden Berechnung der Koreferenzen erläutert und algorithmisch beschrieben. Dabei wird deutlich gemacht, dass es sinnvoll ist, nur Teilgraphen des I-Graphen für den eigentlichen Vergleich von P-Graphen mit dem I-Graphen zu verwenden. Wie diese Teilgraphen berechnet werden können, wird in Abschnitt 6.2 erläutert. Abschnitt 6.3 geht dann auf das Einfügen der Koreferenzkanten und die Bestimmung der Konfidenzwerte ein. Als Abschluss des Kapitels werden in Abschnitt 6.4 die unterbreiteten Vorschläge anhand eines Beispiels veranschaulicht.

6.1 Berechnung der Koreferenzen während der Navigation

Um den Aufbau des folgenden Abschnitts zu erklären, wird zunächst der prinzipielle Ablauf der Wahrnehmung und Koreferenz während der Navigation des Geometrischen Agenten dargestellt. Angenommen, der Geometrische Agent steht an oder kurz vor einem EP (an dem er noch nicht gewesen ist, was im Fall des GA immer zutrifft) und möchte seinen Weg mit Hilfe der Routeninstruktion fortsetzen. Ein rationales Verhalten ist zunächst herauszufinden, welcher Teil der erhaltenen Routeninstruktion dem zuletzt in der Umgebung identifizierten Teil folgt, also dem Teil folgt, für den zuletzt die Koreferenz bestimmt wurde. Ist dieser Teil in der Repräsentation der Routeninstruktion gefunden, so ist es rational zu versuchen, diesen Teil in der

Umgebung zu identifizieren und, wenn eine Koreferenz hergestellt werden konnte, dem in der Umgebung identifizierten Weg zu folgen.

Das bedeutet im Fall des Geometrische Agent, dass im I-Graphen die nächste Kurslandmarke gefunden werden muss. Dann muss geprüft werden, anhand welcher Dinge sich die Kurslandmarke in der Umgebung identifizieren lässt. Dafür ist zuerst zu prüfen, ob die Kurslandmarke in Relation zu anderen Objekte, also Landmarken, lokalisiert ist oder ob die Kurslandmarke als eine Fortsetzung des bisherigen Pfades anzusehen ist, wie es zum Beispiel bei einem EPZD der Fall sein kann. Ist letzteres der Fall, so muss in der Umgebung nur nach einem Weg gesucht werden, der einer guten Fortsetzung des bisherigen Wegs entspricht. Solch eine gute Fortsetzung kann zum Beispiel der Weg sein, der von einer Kreuzung abgeht und die geringste Winkelabweichung vom dem Weg aufweist, auf dem der Agent die Kreuzung erreicht hat.

Sonst muss die eben bereits angesprochenen Identifizierung der Landmarken in der Umgebung vorgenommen werden. Ein Beispiel soll dies verdeutlichen. Angenommen, der Geometrische Agent befindet sich kurz vor der Stelle, an der Beispiel (4.7) anknüpft. Er ist also kurz vor der Region, die zwischen Haus B und Haus C liegt. Er hat die nächste Kurslandmarke im I-Graphen identifiziert. Nun muss er überprüfen, wie diese Kurslandmarke in der Umgebung identifiziert werden kann. Dabei stellt er fest, dass die Kurslandmarke einen Innenpunkt hat, der in einer Region lokalisiert ist, welche durch die räumliche Relation *zwischen* charakterisiert ist und als Referenzobjekte zwei Landmarken hat. Daraufhin erscheint es vernünftig, wenn der Geometrische Agent versucht, in der Umgebung die beiden Landmarken also Haus B und C zu identifizieren. Da er sich in unmittelbarer Nähe der Landmarken befindet, sollte dies kein Problem sein. Sind diese gefunden, so kann er die Region zwischen den beiden Häusern berechnen und versuchen, einen Weg in der Umgebung zu finden, der durch diese Region verläuft.

Angenommen der GA kann, aus welchen Gründen auch immer, die Landmarken oder eine der Landmarken nicht identifizieren, zum Beispiel da es in dem P-Graphen keine Knoten gibt, die Landmarken repräsentieren und die eine hinreichend hohe Ähnlichkeit zu den I-Knoten aufweisen, die die Landmarken der Routeninstruktion repräsentieren. Nun erscheint es sicherlich vernünftig, einen größeren Ausschnitt der Wegbeschreibung zu betrachten und zu versuchen, eine möglichst ähnliche Konstellation in der Umgebung zu identifizieren. Im angenommenen Beispiel ist eine solche Konstellation gegeben durch zwei Häuser, zwischen denen ein Weg hindurch führt. Eine rationale Handlungsmöglichkeit für einen Agenten wäre dann, in der Umgebung nach zwei Häusern zu suchen, zwischen denen ein Weg hindurch führt. Im Geometrischen Agenten kann also der Teilgraph des I-Graphen identifiziert werden, der die eben beschriebene Zwischen-Konstellation repräsentiert und damit das Perzeptionsmodul angewiesen werden, nach Objekten in der Umgebung zu suchen, zwischen denen ein Weg liegt. Der so entstehende P-Graph kann dann mit dem I-Teilgraphen verglichen werden.

Wenn bei dem Vergleich eine Ähnlichkeit berechnet wird, die größer als ein bestimmter Wert ist, kann der Agent annehmen, dass die Konstellation des Teilgraphen in der Umgebung identifiziert wurde.

Diese Steuerung der Aufmerksamkeit durch vorhandenes Wissen entspricht Theorien der Kognitionspsychologie, wie zum Beispiel dem von Neisser in [Nei79] vorgestellten Wahrnehmungszyklus. Der Wahrnehmungszyklus beschreibt die Wahrnehmung als einen nie endenden Kreislauf, der so genannte Schemata für die Steuerung der Aufmerksamkeit verwendet, dadurch (nur) bestimmte Objekte wahrnimmt und durch die neue Information die lenkenden Schemata verändert. Wahrnehmung ist ihrem Wesen nach also selektiv.⁴³ Die ungelenkte Perzeption, wahrzunehmen was wahrzunehmen ist, stellt die Ausnahme dar. Wahrnehmung hängt im Normalfall von spezifischen Zielen ab.⁴⁴ Im Falle des Geometrischen Agenten hängt die Wahrnehmung von dem Ziel ab, seinen Weg mittels der gegebenen Wegbeschreibung zu finden.

Der eben erläuterte Ablauf der Wahrnehmung und Koreferenzierung im Geometrischen Agenten ist durch Algorithmus 6.1.1 beschrieben. Im ersten Schritt von Algorithmus 6.1.1 wird eine Liste mit Teilgraphen des I-Graphen erstellt. Jeder Teilgraph repräsentiert die Lokalisierung einer Position der Kurslandmarke, die zu einem bestimmten Zeitpunkt der Navigation grade relevant ist. Da jede Kurslandmarke des I-Graphen mehrere Positionen enthalten kann und es möglich sein sollte, mehrere Entscheidungspunkte (und somit auch Kurslandmarke) in die Suche mit einzubeziehen, wird eine Liste von Teilgraphen benötigt. Das Einbeziehen mehrerer Entscheidungspunkte in die Suche entspricht der Suchtiefe des Algorithmus. Die Prozedur `SUCHE TEILGRAPHEN`, die die Teilgraphen in Abhängigkeit von der Suchtiefe und einem I-Graphen berechnet, wird in Abschnitt 6.2 erläutert. In Schritt 2 des Algorithmus wird ein P-Graph erzeugt, der die Repräsentationen aller wahrgenommenen Objekte enthält, also CRIL-Knoten mit den wahrgenommenen Attributen der Objekte.⁴⁵

In der Schleife, die in Schritt 3 von Algorithmus 6.1.1 beginnt, wird versucht, für jeden Teilgraph die durch ihn repräsentierte räumliche Konstellation in der Umgebung zu identifizieren. Dabei wird folgendermaßen vorgegangen: Zunächst wird versucht, die räumliche Relation im Teilgraphen zu identifizieren, durch die die Region charakterisiert ist (Schritt 4). In dem oben angegebenen Beispiel würde dabei die Zwischen-Relation bestimmt werden. Wenn es keine räumliche Relation gibt, die eine Region charakterisiert (Schritt 5), so muss der Pfad eine Fortsetzung des vorhergehenden sein. In diesem Fall

⁴³ Vgl. [Eng90] Abschnitt 2.3 und [Nei79] Kapitel 2, 4 und 5.

⁴⁴ Engelkamp in [Eng90] weist darauf hin, dass sogar motorische Aspekte bei der Wahrnehmung eine Rolle spielen, also ob und wie sich der Wahrnehmende bewegt. Solche Aspekte seien hier nicht berücksichtigt. Die Zielorientierung und Selektivität der Wahrnehmung ist aber von großer Bedeutung.

⁴⁵ Dass in Schritt 2 des Algorithmus die Repräsentationen aller wahrgenommenen Objekte erstellt werden, ist sicher nicht als kognitiv adäquat anzusehen. Es wäre allerdings auch kein Problem, die benötigten Knoten erst dann zu erzeugen, wenn sie benötigt werden, also in den Schritten 6, 11 und 18 von Algorithmus 6.1.1.

wird das Perzeptionsmodul angewiesen, nach einem Weg der Umgebung zu suchen, der eine gute Fortsetzung des Weges ist, der vom Agenten zuletzt gegangen wurde. Diese Anweisung an die Perzeption wird durch die Prozedur BESTEFORTS in Schritt 6 gegeben, wobei der P-Graph übergeben wird, in dem ja bereits eine Repräsentation des gesuchten Weges enthalten sein muss. Wie diese Suche auszusehen hat, ist nicht Gegenstand dieser Arbeit. Wenn ein solcher Weg in der Umgebung identifiziert wurde, wird die Koreferenz zwischen der Repräsentation des Weges im P-Graph und im I-Graphen durch die Prozedur FÜGEKOREFEIN in Schritt 6 eingetragen. Die in Schritt 7 verwendete Anweisung **continue** bewirkt, dass der nächste Schleifendurchlauf beginnt, ohne die folgenden Anweisungen des aktuellen Schleifendurchlaufs auszuführen.

Wenn eine räumliche Relation gefunden wurde, werden die Referenzobjekte der räumlichen Relation im Teilgraphen identifiziert. Die Prozedur SUCHEREFOBJ übernimmt diese Suche nach den Referenzobjekten (Schritt 9). Sie gibt eine Liste mit den I-Knoten zurück, die die Referenzobjekte repräsentieren. Im oben angegebenen Beispiel sind die Referenzobjekte zum Beispiel die Knoten, die die Häuser B und C repräsentieren. Der Agent sollte nun versuchen, genau diese Objekte in seiner Umgebung zu identifizieren. Dafür wird mit der Prozedur SUCHEREFOBJ im P-Graphen nach den Knoten gesucht, die die Referenzobjekte repräsentieren. Die Prozedur KOREF sucht im P-Graphen nach Knoten, die eine sehr hohe Ähnlichkeit mit den Knoten des Teilgraphen aufweisen, die die Referenzobjekte repräsentieren (Schritt 10). Sie gibt nil zurück, wenn nicht für alle Referenzobjekte ein Knoten im P-Graphen gefunden wurde, für die eine hohe Ähnlichkeit berechnet wurde.

Werden diese Objekte gefunden, so wird das Perzeptionsmodul angewiesen, die durch die Relation und die Referenzobjekte repräsentierte Region in der Umgebung zu berechnen und die Objekte der Umgebung zu bestimmen, die in dieser Region liegen (Prozedur PERZIPIERERELATIONEN). Für alle Objekte der Umgebung, die in der Region liegen, wird dabei eine entsprechende Repräsentation im P-Graphen eingetragen (Zuweisung in Schritt 11). Das bedeutet, dass im P-Graph ein Knoten erzeugt wird, der von der Sorte Region ist, sowie eine Kante zwischen diesem Knoten und den Referenzobjekten, die die entsprechende räumliche Relation repräsentiert. Für alle Objekte in der Umgebung, die in der Region liegen, wird eine Kante zwischen dem Knoten im P-Graphen, der die Objekte repräsentiert, und dem Knoten, der die Region repräsentiert, erzeugt. Diese Kante repräsentiert wiederum, dass das Objekt in dieser Region lokalisiert ist.

Der so erzeugte P-Graph kann nun mit dem Teilgraph des I-Graphen verknüpft werden, wie es in Kapitel 4 dargestellt wurde (Schritt 12). Ist nun die Ähnlichkeit bezüglich der Verknüpfung größer als ein bestimmter Schwellenwert (Schritt 13), so ist davon auszugehen, dass die erzeugte Verknüpfung der Koreferenz entspricht. Die Prozedur FÜGEKOREFEIN fügt die Koreferenten Knoten und Kanten des P-Graphen sowie die entsprechenden Koreferenzkanten in den I-Graphen ein (Schritt 14).

Ist die Ähnlichkeit unter dem Schwellenwert oder ist es gar nicht möglich gewesen, die Referenzobjekte eindeutig in der Umgebung zu identifizieren, so wird das Perzeptionsmodul angewiesen, alle Objekte der Umgebung als Referenzobjekte zu verwenden (Schritt 18). So würde im oben angegebenen Beispiel zwischen je zwei wahrgenommenen Objekten eine Region zwischen diesen Objekten berechnet werden. Dann prüft die Perzeption, ob es Objekte gibt, die in der Region liegen. Für jedes Objekte, das in der Region liegt, wird im P-Graph eine Kante zwischen dem Knoten, der das Objekt repräsentiert, und dem die Region repräsentierenden Knoten eingefügt. Diese Kanten repräsentieren, dass das Objekt in der Region lokalisiert ist, die Kantenfunktion bildet die Kanten also auf die Lokalisiert-Relation ab. Nun wird der so entstandene P-Graph wieder mit dem I-Teilgraph verknüpft, die Ähnlichkeit bestimmt und, wenn die Ähnlichkeit größer als der Schwellenwert ist, die entsprechenden Knoten und Kanten dem I-Graph zugefügt (Schritte 19-21).

Algorithmus 6.1.1: Gegeben sei ein I-Graph IGraph und eine natürliche Zahl suchtiefe. Es wird ein P-Graph berechnet, der die wahrgenommenen Objekte des GA repräsentiert, sowie evtl. Koreferenzkanten zwischen I-Graph und P-Graph.

Datenstrukturen:

CRIL-Graphen: PGraph, teilGraph
 Verknüpfungen: VK
 Listen: TeilgraphListe, RefObjekte
 String: raumRel

Prozedur:

```

FINDEWEG(IGraph, suchtiefe)46
1. TeilgraphListe := SUCHETEILGRAPHEN ( IGraph, suchtiefe )
2. PGraph := ERSTELLEPKNOTEN(PERZIPIEREOBJ())
3. for all teilGraph ∈ TeilgraphListe do
4.     raumRel := SUCHERAUMREL(teilGraph)
5.     if raumRel = nil do
6.         FÜGEKOREFEIN(IGraph, BESTEFORTS(PGraph))
7.         continue
8.     end else do
9.         RefObjekte := SUCHEREFOBJ(teilGraph)
10.        if KOREF( PGraph , RefObjekte) ≠ nil do
11.            PGraph ← PERZIPIERERELATIONEN(PGraph, raumRel,
                KOREF(PGraph , RefObjekte))
12.            VK := VERKNÜPFEN(PGraph , teilGraph)
13.            if ähnlichkeit(PGraph , teilGraph, VK) > schwelle do
14.                FÜGEKOREFEIN(IGraph, VK)

```

⁴⁶ Es versteht sich, dass viele der Variablen durch Substitution der entsprechenden Methoden ersetzt werden könnten. Wegen der bessern Übersichtlichkeit wurde darauf verzichtet.

```

15.           continue
16.           end
17.       do
18.           PGraph:= PERZIPIERERELATIONEN(PGraph, raumRel)
19.           VK     := VERKNÜPFEN(teilGraph , PGraph)
20.           if ähnlichkeit(PGraph , teilGraph, VK) > schwelle do
21.               FÜGEKOREFEIN(IGraph, VK)
22.           end
23.       end
24.   end
25. end

```

6.2 Berechnung der relevanten Teilen des Instruktionsgraphen

In diesem Abschnitt soll nun gezeigt werden, welche Elemente eines gegebenen I-Graphen ein I-Teilgraph enthalten muss, wie er in Algorithmus 6.1.1 benötigt wird. Es lassen sich sofort zwei Feststellungen treffen. Zum einen ist es nicht nötig die Knoten im I-Graph, die bereits in der Koreferenzrelation stehen, noch einmal in die Suche mit einzubeziehen. Zum anderen ist es nicht sinnvoll, Knoten zu betrachten, die bezüglich des momentanen Standorts des Agenten weit entfernt liegen. Die Frage, wie viel der Wegbeschreibung der Agent zu einem Zeitpunkt der Navigation versuchen sollte, in der Umgebung zu identifizieren, soll und kann nicht Gegenstand der vorliegenden Arbeit sein. Die Beantwortung der Frage hängt sicherlich von vielen Faktoren ab, wie z.B. dem geologischen Gegebenheiten der Umgebung (in einer verwinkelten Altstadt macht es keinen Sinn, nach Entscheidungspunkten zu suchen die weit entfernt liegen), der Persönlichkeit des Instruktors und vielem mehr. Es ist m. E. aber klar, dass nicht der ganze ‚Rest‘ der Route zu jedem beliebigen Zeitpunkt der Navigation relevant ist.

Zunächst soll geklärt werden, wie die Struktur der I-Teilgraphen prinzipiell aussieht, also welche Element in einen Teilgraphen enthalten sein müssen. Das wichtigste Element in einem Teilgraphen ist der Pfad, da er den eigentlichen Weg der Route repräsentiert. Der Pfad wird durch seine räumliche Lage in Bezug auf Landmarken, seien es nun Plätze oder andere Objekte der Umgebung lokalisiert. Es lässt sich also folgendes festhalten:

- Landmarken dienen der Charakterisierung von Regionen
- Regionen dienen der Lokalisierung von Position
- Positionen dienen der Charakterisierung von Pfaden.

Angenommen, der Agent ist an einem Entscheidungspunkt angekommen. Der nächste interessante Teil der Routeninstruktion beschreibt den oder die Pfade, die zwischen seinem momentanen Standort und dem nächsten EP liegen. Diese Pfade sind nach (5.4.1) unter einer Kurslandmarke zusammengefasst. Informell müssen die I-Teilgraphen also mindestens aus folgenden Elementen bestehen:

- Eine Kurslandmarke
- Eine Menge von Pfaden, die in der Kurslandmarke enthalten sind
- Eine Menge von Positionen, die die Pfade und die Kurslandmarke charakterisieren
- Eine Menge von Regionen, für die gilt, dass sie entweder
 - einen Pfad als Referenzobjekt haben oder
 - ein Pfad bzw. eine der Positionen des Pfades in ihnen enthalten ist

Für eine gezielte Suche nach bestimmten Dingen in der Umgebung ist ein solcher I-Teilgraph aber immer noch sehr ungünstig, da er sehr groß sein kann. Es ist also vorteilhaft, den I-Teilgraph noch weiter zu unterteilen. Eine Möglichkeit, dies zu tun, ist jede durch den I-Teilgraph repräsentierte Objektlokalisierung wiederum durch einen Teilgraphen des I-Teilgraphen zu repräsentieren. Dieser Teilgraph des I-Teilgraphen soll im Folgenden I_{REG} -Teilgraph heißen. Dabei sind die Regionen die zentralen Elemente. Sie stellen die Verbindung zwischen den Landmarken und den Pfaden her. Die I_{REG} -Teilgraphen haben dann (informell) folgende Struktur:

- Eine Region.
- Eine oder mehrere Landmarken, welche die Referenzobjekte sind, durch die die Region charakterisiert ist.
- Alle Knoten die in der lokalisiert-Relation zu der Region stehen.

Die Struktur der I_{REG} -Teilgraphen hat folgenden Vorteile:

- Die Größe der Teilgraphen ist (relativ) gering.
- Es sind nur genau die Landmarken in den Teilgraphen enthalten, die Referenzobjekte sind, die also zur Charakterisierung der Region in der Umgebung notwendig sind.

Die so bestimmten I_{REG} -Teilgraphen ermöglichen es, der Perzeption genau mitzuteilen, welche räumlichen Relationen gesucht werden (vgl. Algorithmus 61.1), womit die Anzahl der während der Wahrnehmung zu berechnenden Regionen⁴⁷ relativ klein ist. Außerdem sind die zu vergleichenden Graphen von relativ geringer Größe, der Berechnungsaufwand (nicht aber die Komplexität) wird damit eingeschränkt und die Antwortzeit verkürzt. Dass dieses Vorgehen

⁴⁷ Wie solche Regionen in einer zweidimensionalen Domäne wie der Domäne des GA berechnet werden können, wird in [Sch99] und [Hel02] beschrieben.

auch als kognitionspsychologisch Adäquatheit angesehen werden kann, wurde bereits im vorigen Abschnitt mit der von natur aus selektiven Wahrnehmung begründet. Im Folgenden wird ein Vorschlag gemacht, wie die eben eingeführten Teilgraphen in einem gegebenen I-Graphen gefunden werden können. Zuerst wird dafür der Algorithmus 6.2.1 zum Finden des I-Teilgraphen in einem gegebenen I-Graphen beschrieben. Dann wird die weitere Einschränkung auf die I_{REG} -Teilgraphen informell beschrieben.

In Schritt 1 von Algorithmus 6.2.1 wird der Knoten des I-Graphen bestimmt, der den momentanen Standort des GA repräsentiert. Der Standort kann entweder ein Pfad sein, wenn der Agent sich grade auf der Kurslandmarke befindet, die den Pfad enthält, oder eine Position, wenn der Agent sich an dem Platz befindet, der durch die Position charakterisiert ist. Wie der Knoten, der den Standort repräsentiert, bestimmt wird, soll an dieser Stelle nicht beschrieben werden. Ist der Knoten ein Pfad, so wird der Endpunkt des Pfades als Startzustand der Suche verwendet (Zeile 2). Der Startzustand ist folglich immer eine Position.

Nun werden alle vom Startzustand abgehenden Pfade gesucht (Zeile 5). Abgehend bedeutet, dass der Startzustand der Startpunkt des Pfades ist. Danach werden für alle abgehenden Pfade alle Innenpunkte gesucht (Zeile 7). Für alle so ermittelten Positionen wird wiederum die Menge aller Regionen berechnet, in denen sie lokalisiert oder explizit nicht lokalisiert sind (Zeile 9). Die Menge der so bestimmten Regionen wird zur Liste der Regionen hinzugefügt. Abschließend wird der Endpunkt von jedem Pfad bestimmt, sowie für diesen Endpunkt die Menge der Regionen, in denen er lokalisiert oder explizit nicht lokalisiert ist. Diese Menge wird ebenfalls der Liste der Regionen hinzugefügt. Ist der Endpunkt ein Entscheidungspunkt, so wird der Zähler der Schleife um 1 inkrementiert. Anschließend wird der Endpunkt zum Startzustand und die Schleife läuft erneut durch.

Ist die in Zeile 4 beginnende Schleife beendet, sind die Regionen, die den oder die nächsten Pfadabschnitte identifizieren, bestimmt und in der Liste der Regionen enthalten. Für jede so bestimmte Region wird in Zeile 17 durch die Prozedur `BERECHNEREGIONSTEILGRAPH` ein Teilgraph berechnet, der wiederum der Ausgabeliste hinzugefügt wird.⁴⁸

Algorithmus 6.2.1: *Gegeben sei ein I-Graph $I\text{Graph}$ und eine natürliche Zahl suchtiefe. Es wird eine Liste von Regionsteilgraphen des I-Graphen berechnet.*

Datenstrukturen:

CRIL-Knoten:	start, ck, ck2
CRIL-Graphen:	ausgabeGraph
Listen:	L, regListe, ausgabeListe
Integer:	i, j

⁴⁸ Es bietet sich an, für die Ausgabeliste immer Paare, bestehend aus einer Region und dem durch sie bestimmten Teilgraphen zu verwenden. So wird eine erneute Suche nach der Region des Teilnetzes in Algorithmus 5.1.1 vermieden.

Prozedur:

```
SUCHETEILGRAPHEN(IGraph, suchtiefe)
1. start := AGENTENSTANDORT(iGraph)
2. if (Pfad ∈  $\mathcal{S}$ (start)) start := ENDPUNKT(start)
3. i := 0
4. while ( i < suchtiefe ) do
5.     pfade := ABGEHENDEPFADE(start)
6.     for all (ck2 ∈ pfade) do
7.         positionen := INNENPUNKTE(ck2)
8.         for all (ck3 ∈ positionen) do
9.             regionen := REGIONEN(ck3)
10.            FÜGEEIN(regListe, regionen)
11.        end
12.        FÜGEEIN(regListe, REGIONEN(ENDPUNKT(ck2)))
13.        if (ISTEP(ENDPUNKT(ck2))) i := i + 1
14.        start := ENDPUNKT(ck2)
15.    end
16. end
17. for all ( ck ∈ regListe ) do
18.     ausgabeGraph := REGIONSTEILGRAPH(ck , iGraph)
19.     FÜGEZU(ausgabeListe, ausgabeGraph)
20. end
21. return ausgabeListe
```

Die Prozedur BERECHNEREGIONSTEILGRAPH bestimmt für eine Region eines I-Graphen einen Teilgraphen des I-Graphen, der folgende Elemente enthält:

- Die Referenzobjekte der Region. Handelt es sich um einen Pfad, dann auch seine Start- und Endpunkte und die Kurslandmarke, in der der Pfad enthalten ist, sowie alle Kanten zwischen Pfad, Positionen und Kurslandmarke
- Alle Knoten, die durch die Region lokalisiert sind.
- Alle Kurslandmarke, die einen beliebigen Punkt haben, der durch die Region lokalisiert ist, sowie alle enthaltenen Pfade.
- alle Plätze, die eine beliebige Position haben, der durch die Region lokalisiert ist.

Ein Punkt bezüglich Plätzen wurde bisher noch nicht berücksichtigt und soll an dieser Stelle nur kurz erläutert werden. Die Charakterisierung von Plätzen in Routeninstruktionen erfolgt nicht unbedingt nur über die Lokalisierung von Positionen des Platzes. Es werden auch Ausdrücke verwendet, wie zum Beispiel „... von da siehst du ein hohes Haus...“. In diesem Beispiel wird also nur ausgedrückt, dass von der Position des Platzes aus ein bestimmtes Objekt sichtbar ist. Das würde nicht zu einer Repräsentation im I-Graphen führen, in der die Position in einer Region lokalisiert ist. Wie solche Ausdrücke einer

Routeninstruktion in CRIL-Graphen repräsentiert werden, steht aber zum jetzigen Zeitpunkt noch nicht fest. Es sollte aber ohne weiteres möglich sein, auch für solche Ausdrücke Teilgraphen zu bestimmen. Dafür könnte in Algorithmus 6.2.1 für jede betrachtete Position geprüft werden, ob eine entsprechende Repräsentation im I-Graphen enthalten ist. Ist dies der Fall, wird die Position zusammen mit der Repräsentation in einer weiteren Liste mit Teilgraphen gespeichert, die dann zusammen mit der Liste der Regionsteilgraphen von Algorithmus 6.2.1 zurückgegeben wird. Sind solche Teilgraphen vorhanden, kann das Perzeptionsmodul angewiesen werden, nach eben solchen Konstellationen in der Umgebung zu suchen und anschließend kann ein Vergleich durchgeführt werden, der analog zu dem der Regionsteilgraphen verläuft.

6.3 Einfügen der Koreferenzkanten und Konfidenzwertberechnung

In diesem Abschnitt soll informell erläutert werden, wie die Prozedur FÜGEKOREFEIN aus Algorithmus 6.1.1 die Knoten des P-Graphen aus einer Verknüpfung in den I-Graphen einfügt und anschließend die Koreferenzkanten zwischen den P-Knoten und den I-Knoten erzeugt. Die Koreferenzkante soll die extensionale Gleichheit zweier Knoten aus verschiedenen Kontexten repräsentieren. Zwar soll sie dies auch im vorliegenden Fall, es muss aber klar sein, dass es sich eher um eine vermutete Koreferenz handelt. Je höher die in Algorithmus 6.1.1 bestimmte Ähnlichkeit zweier eines I_{REG} -Teilgraphen mit einem P-Graphen, desto größer die Wahrscheinlichkeit, dass es sich bei den wahrgenommenen Objekten um die in der Routeninstruktion beschriebenen Objekte handelt.

Ist nun in Algorithmus 6.1.1 die Ähnlichkeit der verglichenen Graphen größer als der gegebene Schwellenwert, so wird die Prozedur FÜGEKOREFEIN aufgerufen. Das eigentliche Einfügen der Koreferenzkanten ist trivial. Jedes Knotenpaar der Verknüpfung repräsentiert eine Koreferenzkante. Zunächst werden die Knoten und Kanten der übergebenen Verknüpfung in den übergebenen I-Graphen eingefügt, die nicht schon im I-Graphen enthalten sind. Dies müssen nämlich grade die Knoten aus dem P-Graphen sein, die in der Verknüpfung enthalten sind. Dann werden entsprechend der Knotenpaare die Koreferenzkanten zwischen den eingefügten P-Knoten und den in der Verknüpfung enthaltenen I-Knoten eingefügt.

Der Konfidenzwert der Koreferenzkanten, der die Sicherheit des Agenten bezüglich der Koreferenzen repräsentiert, kann auf verschiedene Weisen bestimmt werden. Zum einen ist es möglich, die Ähnlichkeiten von Knoten als Grundlage der Berechnung für den Konfidenzwert der Koreferenz zwischen je zwei Knoten zu verwenden. Damit wird für jedes Knotenpaar ein eigener Konfidenzwert berechnet und somit jeder Koreferenzkante ein eigener Wert zugeordnet. Die andere Möglichkeit besteht darin, den Wert der Ähnlichkeit

ganzer I-Teilnetze als Grundlage zu verwenden. Dies würde bedeuten, dass alle Koreferenzkanten eines Teilnetzes den gleichen Konfidenzwert zugeordnet bekommen, der sich auf die Ähnlichkeit bezüglich der an FÜGEKOREFEIN übergebenen Verknüpfung bezieht.

Welches der beiden Verfahren sinnvoller ist, lässt sich zum jetzigen Zeitpunkt noch nicht sagen. Für die Verwendung der Knotenähnlichkeit spricht, dass die Koreferenzkante letztendlich für die extensionale Gleichheit der beiden in Relation stehenden Knoten steht. Andererseits steht die Ähnlichkeit der Verknüpfung der ganzen Teilgraphen für die Güte des Vergleichs. Somit lassen sich aus ihr auch Rückschlüsse über die Sicherheit der vorgenommenen Koreferenz ziehen. Wenn der Geometrische Agent vollständig implementiert ist, lässt sich mit ihm prüfen, welches der Verfahren zu einer besseren Navigationsleistung des Agenten führt.

6.4 Ein Beispiel

In diesem Abschnitt soll an einem kleinen Beispiel der prinzipielle Ablauf der Integration von Instruktion und Perzeption verdeutlicht werden. Die Routeninstruktion, die für dieses Beispiel verwendet wird, entstammt [TSH03] und beschreibt einen Weg von der Mensa zu Haus E auf dem Campus des Fachbereichs Informatik in Hamburg.

(6.1) „Wenn du aus der Mensa kommst, geh nach links, zwischen Haus B und Haus C durch. Geh hinter Haus C lang, und dann, wenn du an Haus C vorbei bist, wieder nach rechts. Dann stehst du vor Haus E.“

Dass es sich bei dieser Wegbeschreibung um eine nach menschlichem Verständnis gute Wegbeschreibung handelt, lässt sich anhand Abbildung 2.4 auf S. 27 nachvollziehen, die eine Karte des Campus zeigt. Die Mensa des Fachbereichs ist in Haus B. Ein vereinfachter, den ersten Teil der Routeninstruktion repräsentierender I-Graph, der in der Instruktionsphase berechnet wurde, ist in Abbildung 6.1 dargestellt, die Taxonomie der Eigenschaften in Abbildung 5.1 (S. 77). Der GA soll bereits an der Abzweigung vor der Mensa angekommen sein. Die erste Koreferenzierung wurde also bereits durchgeführt, was in Abbildung 6.1 durch den gekennzeichneten P-Graphen und die rot gestrichelte Koreferenzkanten dargestellt ist.

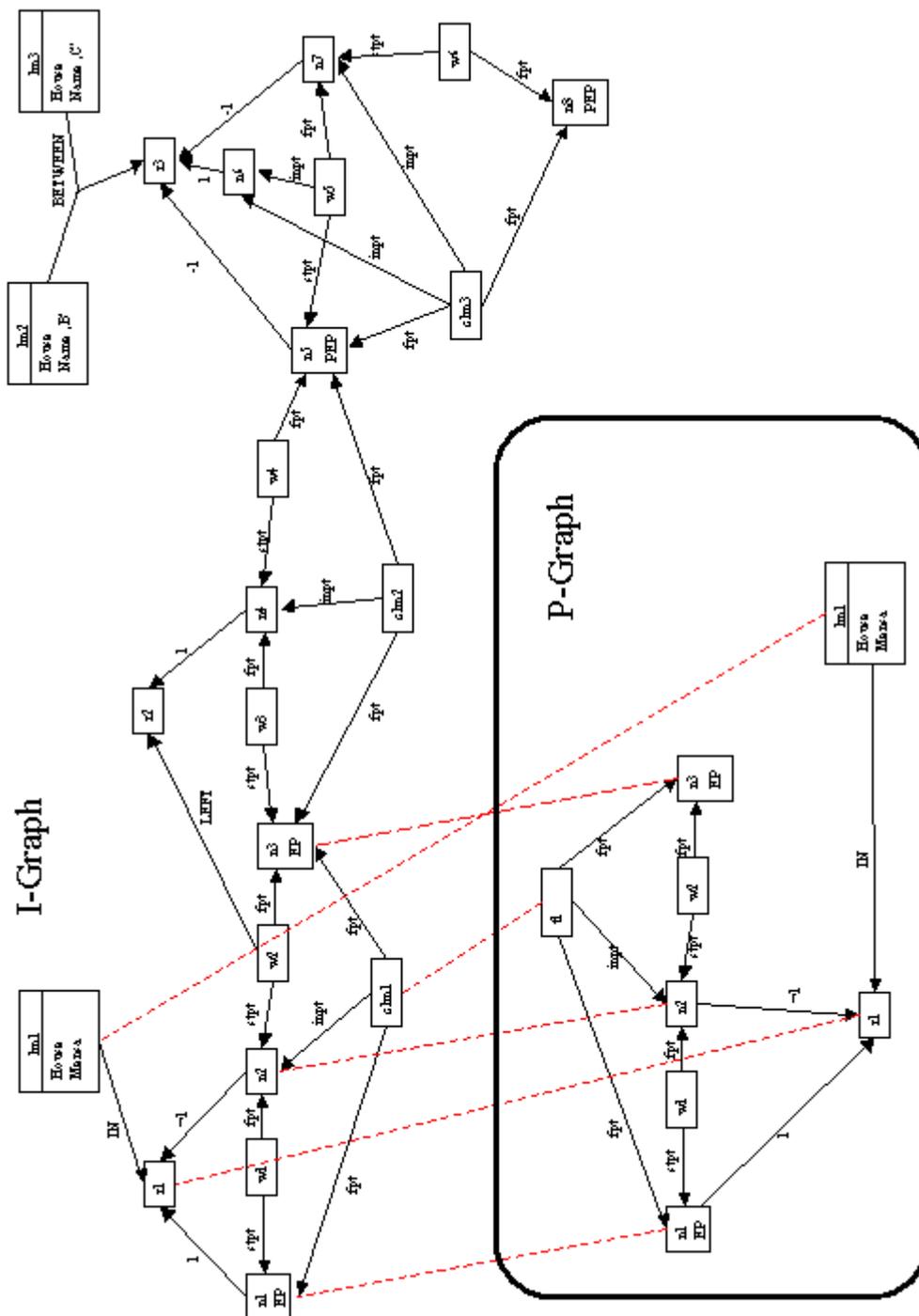


Abbildung 6.1: Ein (vereinfachtes) Beispiel (6.1) repräsentierender I-Graph und der erste P-Graph, für den die Koreferenz bereits ermittelt wurde. Die roten, gestrichelten Linien stellen Koreferenzkanten dar.

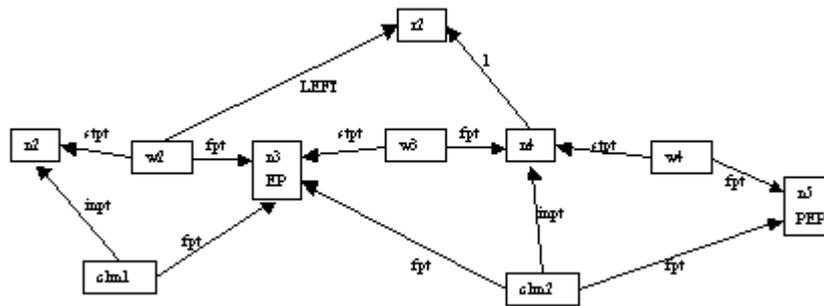


Abbildung 6.2: Teilgraph des I-Graphen der die Lokalisierung des nächsten Routenteils repräsentiert.

Um nun den nächsten Teil der Route zu identifizieren, wird vom Planungsmodul Algorithmus 6.1.1 ausgeführt. Algorithmus 6.1.1 führt im ersten Schritt Algorithmus 6.2.1 aus. Die Variable 'suchtiefe' soll 1 sein. Im ersten Schritt von Algorithmus 6.2.1 wird dann im I-Graphen der momentane Standort des Agenten ermittelt. Da hier angenommen wird, dass der Agent sich bereits an dem Entscheidungspunkt vor der Mensa befindet, wird der Knoten mit der Bezeichnung n3, als der Knoten des momentanen Standorts ermittelt. Der einzige abgehende Pfad von n3 ist w3, was in Schritt 5 berechnet wird. Innenpunkte hat der Pfad nicht, dafür aber einen Endpunkt. Der Endpunkt von w3 ist auch in einer Region lokalisiert, nämlich r2. So wird in Schritt 12 des Algorithmus r2 in die Liste der Regionen aufgenommen. Da der Endpunkt von w3 kein EP ist, wird der Zähler der Schleife in Schritt 13 nicht inkrementiert, was zur Folge hat, dass ein weiterer Schleifendurchlauf erfolgt, in dem der neue Startknoten der Endpunkt von w3 ist, also n4. Beim erneuten Schleifendurchlauf werden keine weiteren Regionen ermittelt, aber die Schleife terminiert, da n5 ein möglicher EP ist.

In der folgenden Schleife in Algorithmus 6.2.1 wird für die Region r2 der

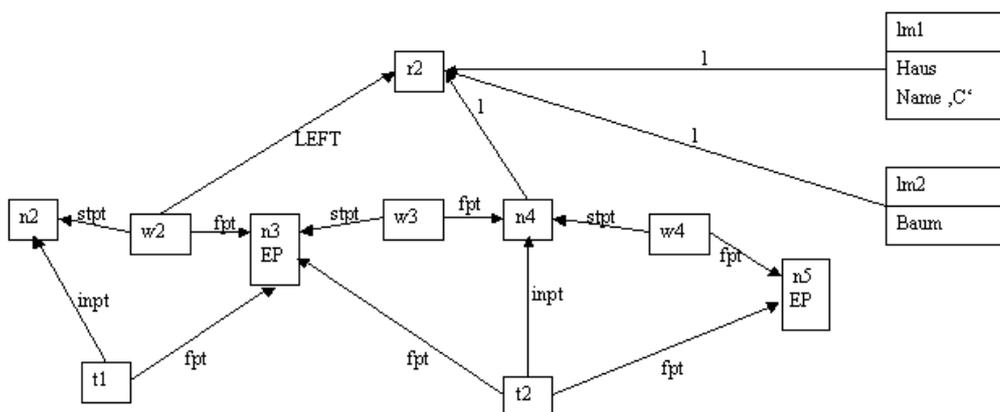


Abbildung 6.3: Der mittels des I-Teilgraphen aus Abbildung 6.2 von der Perzeption des Agenten erstellte P-Graph.

von Knoten oder Kanten zum Beispiel das zweite Element des Paares vertauscht wird (nur so ist die Konsistenz gewahrt), oder bei einer Verknüpfung das Argument der Paares, das aus dem P-Graphen stammt, gegen ein nicht verknüpftes Element getauscht wird. So wird sich immer eine kleinere Ähnlichkeit ergeben. Zu beachten ist auch, dass die Ähnlichkeit nicht abnimmt, wenn die Knoten des P-Graphen mehr Eigenschaften haben, also vom Perzeptionsmodul mehr Eigenschaften wahrgenommen wurden. Dies ergibt sich direkt aus Definition 4.3.3 für die Bestimmung der Ähnlichkeit von Knoten. Würde man, wie am Ende von Abschnitt 4.2 vorgeschlagen wurde, berücksichtigen, dass die Ähnlichkeit von einem Konzept der Routeninstruktion und einem wahrgenommenen Konzept immer dann sehr groß ist (zum Beispiel 1), wenn das wahrgenommene Konzept vom Instruierten Konzept subsumiert wird, ergibt sich sogar eine Ähnlichkeit von 0,998, da die Sorte Weg von Kurslandmarke subsumiert wird.

Kapitel 7

Ausblick

Da der GA noch in der Entwicklung ist, kann die Frage, ob die in dieser Arbeit vorgestellten Prozesse und Berechnungen zu einer guten Navigationsleistung eines Agenten führen, noch nicht beantwortet werden. Dementsprechend lässt sich die kognitive Adäquatheit der vorgestellten Prozesse im Kontext des Modells eines navigierenden Agenten auch erst feststellen, wenn der GA vollständig implementiert ist. Wenn das System des Geometrischen Agenten implementiert ist, sollten die verschiedenen in dieser Arbeit vorgestellten Ansätze zur Bestimmung der Ähnlichkeit von Konzepten, Knoten und Graphen anhand von Versuchen miteinander verglichen werden. Welcher der Ansätze am besten für die Bestimmung der Koreferenzen zwischen I-Knoten und P-Knoten geeignet ist, lässt sich nur anhand einheitlicher Versuche feststellen, in denen jeweils die verschiedenen Ansätze verwendet werden. Es versteht sich, dass die Ergebnisse der Versuche mit empirischen Untersuchungen mit Testpersonen verglichen werden müssen, um zu bestätigen, dass die verwendeten Ansätze als kognitiv Adäquat angesehen werden können (vgl. Abschnitt 1.1.4).

Einige offene Punkte im Bereich der Integration von Graphen der Perzeption und der Instruktion wurden bereits in den jeweiligen Abschnitten erläutert. Hier sollen nun noch einige Anregungen gegeben werden, welche Aspekte der Integration von Graphen aus der Perzeption und der Instruktion weitere Untersuchungen erfordern. Ein Aspekt betrifft die Größe der Teilgraphen aus der Instruktion, ein anderer Aspekt die Konfidenzwerte der Koreferenzkanten. Die bisher bestimmten Teilgraphen aus der Instruktion sind relativ klein. Eine Möglichkeit, die Sicherheit der bestimmten Koreferenz zu erhöhen, wäre größere Ausschnitte des I-Graphen, die mehrere der in dieser Arbeit vorgestellten Teilnetze vereinen, für den Vergleich zu verwenden. Dabei könnte zunächst wie in Kapitel 6 beschrieben vorgegangen werden und danach die Ähnlichkeit des vereinten I-Teilgraphen mit den koreferenzierten P-Knoten und Kanten bestimmt werden. Wird auch für größeren Ausschnitt des I-Graphen eine hohe Ähnlichkeit festgestellt, so wäre der Schluss gerechtfertigt, dass die Sicherheit der bereits eingefügten Koreferenzkanten sehr groß ist. Zu überlegen wäre weiterhin, inwieweit strukturelle Aspekte, wie zum Beispiel

der Knotengrad, bei der Bestimmung der Ähnlichkeit eine Rolle spielen soll oder kann.

Die Konfidenzwerte der Koreferenzkanten sind bisher die berechneten Ähnlichkeiten für die Verknüpfung der koreferenten Knoten. Ob dies in jedem Fall sinnvoll ist, ist noch nicht klar. Um diese Frage beantworten zu können, muss zunächst klar sein, wie die Konfidenzwerte interpretiert werden. Interpretiert werden die Konfidenzwerte vom Planungsmodul des Geometrischen Agenten. Da das Planungsmodul zum jetzigen Zeitpunkt in der Entwicklung ist, ist noch nicht klar, wie die Konfidenzwerte interpretiert werden. Möglich sind zum Beispiel einfache Unterteilung der Ähnlichkeiten in Intervalle, denen sprachliche Ausdrücke zugeordnet werden. Ein Beispiel dafür ist eine Koreferenz mit dem Konfidenzwert ‚sehr sicher‘. Solche diskreten Konfidenzwerte können auch in einer Fuzzy-Logik verwendet werden, wobei die Abbildung von den Ähnlichkeiten auf die Konfidenzwerte dann auch Überlappungen der Intervalle zulässt. Grundsätzlich wäre es auch denkbar, eine eigene Algebra für Konfidenzwerte zu entwickeln, mit deren Werten dann numerisch weiter gerechnet werden kann und eine wie auch immer geartete Abbildung der Ähnlichkeitswerte auf die Konfidenzwerte zu bestimmen.

Literaturverzeichnis

- [BCM03] Baader, F., Calvanese, D., McGuinness, D., Nardi D., Patel-Schneider, P.: *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, Cambridge, 2003.
- [Ben01] Ben-Ari, M.: *Mathematical Logic for Computer Science*. Springer, London, 2001.
- [BFL83] Brachman, R.J., Fikes, R.E., Levesque, H.J.: Krypton: A functional Approach to knowledge Representation. In: *Computer*, 16(10), S. 67-73.
- [BHN97] Brewka, G., Habel, C. und Nebel B.: *KI-97: Advances in Artificial Intelligence*, Band 1303 der Reihe Lecture Notes in Artificial Intelligence. Springer, Berlin, 1997.
- [Bos86] Bosshardt, H.-G. (Hrsg.): *Perspektiven auf Sprache*. De Gruyter, Berlin, 1986.
- [BPN96] Bloom, P., Peterson, M.A., Nadel, L., Garret, M.F.: *Language and Space*. MIT-Press, Cambridge, 1996.
- [BrL85] Brachman, R.J., Levesque, H.J. (Hrsg.): *Readings in Knowledge Representation*. Morgan Kaufmann, Los Altos, 1985.
- [Bun83] Bundy, A. (Editor): *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83)*. Morgan Kaufmann, Los Altos, 1983.
- [CBK00] Casakin, H., Barkowsky, T., Klippel, A., Freksa, Ch.: Schematic maps as Wayfinding Aids. In: [FBH00] S. 54-71.
- [ChM92] Chein, M., Mugnier, M.-L.: Conceptual Gaps: fundamental notions. In: *Revue d'Intelligence Artificielle*, Vol.6, n° 4, S. 365-406, 1992.
- [CoO02] Coventry, K.R., Olivier, P., *Spatial Language: Cognitive and Computational Aspects*. Dordrecht, Kluwer, 2002.
- [CSC00] Cairo, O., Sucar, L.E., Cantu, F.J. (Hrsg.): *MICAI-2000: Advances in Artificial Intelligence*. Springer, Berlin, 2000.

- [DaD98] Daniel, M.-P., Denis, M.: Spatial descriptions as navigational aids: A cognitive analysis of route directions. *Kognitionswissenschaft* 7, S. 45–52. Springer, Berlin, 1998.
- [Deh03] Dehm, S.: Erkennung von Entscheidungspunkten bei Wegbeschreibungen von urbanen Routen. Studienarbeit am Fachbereich Informatik der Universität Hamburg, 2003.
- [Dic45] Dice, L. R.: Measures of the Amount of Ecologic Associations between Species. In: *Journal of Ecology*, 26, S. 297-302, 1945.
- [EHL02] Eschenbach, C., Habel C., Leßmöllmann, A.: Multiple frames of reference in interpreting complex projective terms. In: [CoO02].
- [ELR95] Ellis, G., Levinson, R., Rich, W. Sowa, J. F. (Hrsg.): *Conceptual Structures: Applications, Implementations and Theory*. Springer, Berlin, 1995.
- [Eng90] Engelkamp, J.: *Das menschliche Gedächtnis*. Verlag für Psychologie, Göttingen, 1990.
- [ETH00] Eschenbach, C., Tschander, L., Habel, Ch.: Lexical Specifications of Paths. In: [FBH00] S. 127-144.
- [Esc00] Eschenbach, C.: Geometric structures of frames of reference and natural language semantics. In: *Computation and Cognition*, 1, S. 329-348, 2000.
- [EsK97] Eschenbach, C., Kulik, L.: An axiomatik approach to the spatial relations underlying ‘left’-‘right’ and ‘in front of’-‘behind’. In: [BHN97], S. 207-218.
- [FBH00] Freska, Ch., Bauer, W. , Habel, Ch., Wender, K. F. (Hrsg.): *Spatial Cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications*. Springer, Berlin/Heidelberg, 2000.
- [FBH03] Freksa, F., Brauer, W., Habel, C., Wender, K.F.: *Spatial Cognition III*. Springer, Heidelberg, 2002.
- [FHW98] Freksa, C., Habel, C., Wender, K.F. (Hrsg.): *Spatial Cognition: An interdisciplinary approach to representing and processing spatial knowledge*. Springer, Berlin, 1998.
- [FrM99] Freksa, Ch., Marks, D.M.: *Spatial information theory: Cognitive and computational foundations of geographic information science*. Springer, Berlin, 1999.
- [GHR83] Günther, S., Habel, C., Rollinger, C.-R.: Ereignissnetze: Zeitnetze und referentielle Netze. KIT-Report 12, Berlin, 1983.

- [GiM97] Gillner, S., Mallot, H.A.: Navigation and Acquisition of spatial knowledge in a virtual maze. Technical Report No. 45, Max-Planck-Institute for Biological Cybernetics, Tübingen, 1997.
- [Gol99] Golledge, R.G. (Hrsg.): *Wayfinding behaviour*. Johnson Hopkins University Press, London, 1999.
- [Güt94] Güting, R.H.: GraphDB : A data model and query language for graphs in databases. Informatik Berichte 155, Hagen, 1994.
- [Hab86] Habel, Ch.: *Prinzipien der Referentialität*. Springer, Berlin, 1986.
- [Hab88] Habel, Ch.: Prozedurale Aspekte der Wegplanung und Wegbeschreibung. In: [ScR88], S. 107-133.
- [Hab89] Habel, Ch.: Zwischen-Bericht. In: [HHR89], S. 37-69.
- [HaS76] Hankamer, J., Sag, I.: Deep and surface anaphora. In: *Linguistic Inquiry* 7, S. 391-428. 1976.
- [Hel02] Helwich, J.H.: *Berechnung von räumlichen Relationen in 2-dimensionalen Domänen*. Studienarbeit am FB Informatik der Universität Hamburg, 2002.
- [Her94] Hernández, D.: *Qualitative Representation of Spatial Knowledge, Lecture Notes in Artificial Intelligence 804*. Springer, Berlin, 1994.
- [HeS98] Hermann, T., Schweizer, K.: *Sprechen über Raum*. Hans Huber, Bern, 1998.
- [HHR89] Habel, Ch., Herweg, M., Rehkämper, K. (Hrsg.): *Raumkonzepte in Verstehensprozessen*. Niemeyer, Tübingen, 1989.
- [Hun35] Huntington, E. V.: Interrelations among the four principle types of order. In: *Transactions of the American Mathematical Society*, 38, S. 1-9, 1935.
- [JPP97] Jarke, M., Pasedach, K., Pohl, K. (Editors): *Informatik '97*. Springer, Berlin, 1997.
- [Kle79] Klein, W.: Wegauskünfte. In: *Zeitschrift für Literaturwissenschaft und Linguistik* 33, S. 9-58, 1979.
- [KMI95] Kasahara, K., Matsuzawa, K., Ishakawa, T., Kawaoka, T.: Viewpoint-Based Measurement of Semantic Similarity between Words. In: *Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics*, S. 292-302, 1995.
- [Kui94] Kuipers, B.: *Qualitative reasoning: Modelling and Simulation with incomplete knowledge*. MIT-Press, London, 1994.
- [LeS92] Lellahi, S.K., Spyratos, N.: A graph-based data model: Query language and semantics. Rapport de Recherche 742, Orsay, 1992.

- [Lev86] Levelt, W.J.M.: Zur sprachlichen Abbildung des Raumes: Deiktische und intrinsische Perspektive. In: [Bos86], S. 187-211.
- [Lev96] Levinson, S.: Frames of reference and Molynoeux's question: Crosslinguistic evidence. In: [BPN96], S. 109-169.
- [Lyn73] Lynch, K.: *The Image of the City*. MIT-Press, Cambridge, 1973.
- [MaG98] Mallot, H.A., Gillner, S.: View-based vs. Place-based navigation: What is recognized in recognition-triggered response? Technical report 64, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany.
- [Mal99] Mallot, H.A.: Spatial Cognition: Behavioral competences, neural mechanisms and evolutionary scaling. In: *Kognitionswissenschaft* 8, S. 40-48, 1999.
- [MGL00a] M. Montez-y-Gomez, A. Gelbukh, A. Lopez-Lopez, R. Baeza-Yates: Flexible Comparison of Conceptual Graphs. In: [MLQ01], S. 102-111.
- [MGL00b] M. Montes-y-Gomez, A. Gelbukh, A. Lpez-Lpez (2000). Comparison of Conceptual Graphs. O. Cairo, In: [CSC00] S. 548-556.
- [MLQ01] Mayr, H.C., Lazansky, J., Quirchmayr, G., Vogel, P. (Hrsg.): *Database and Expert Systems Applications*. Springer, Heidelberg, 2001.
- [Mor81] Morris, R.G.M.: Spatial Localization Does Not Require the Presence of Local Cues. In: *Learning and Motivation* 12, S. 239-260, 1980.
- [Nei79] Neisser, U.: *Kognition und Wirklichkeit: Prinzipien und Implikationen der kognitiven Psychologie*. Klett-Cotta, Stuttgart, 1979.
- [OKe91] O'Keefe, J.: The hippocampal cognitive map and navigational strategies. In: [Pai91] S. 273-295.
- [Pai91] Paillard, J. (Hrsg.): *Brain and Space*. Oxford University Press, Oxford, 1991.
- [PoC95] J. Poole, J.A. Campbell: A Novel Algorithm for Matching Conceptual and Related Graphs. In: [ELR95], S. 293 – 307.
- [Por96] Portugali, J. (Hrsg.): *The construction of cognitive maps*. Kluwer, Dordrecht, 1996.
- [Pou93] Poucet, B.: Spatial Cognitive Maps in Animals. In: *Psychological Review*. Vol.100. No.2, S. 163-182, 1993.
- [Pri02] Priss, U. (Hrsg.): *Conceptual structures: integration and interfaces*. Springer, Berlin, 2002.
- [Pyl80] Pylyshyn, Z.W.: Computation and Cognition: Issues in the Foundation of Cognitive Science. In: *Behavioral and Brain Science* 3. S. 111-169, 1980.

- [Rei91] Reimer, U.: *Einführung in die Wissensrepräsentation*. Teubner, Stuttgart, 1991.
- [Res99] Resnik, Ph.: Semantic Similarity in a Taxonomy: An Information Based Measure and its Application to Problems of Ambiguity in Natural Language. In: *Journal of Artificial Intelligence Research 11*, S. 95-130, 1999.
- [RuN03] Russel, S. , Norvig, P.: *Artificial Intelligence : A modern Approach*. Prentice Hall, New Jersey, 2003.
- [RWS98] Rothkegel, R., Wender, K.F., Schumacher, S.: Judging spatial relations from memory. In: [FW98] S. 79-106.
- [Sch38] Schmidt, A.: Über deduktive Theorien mit mehreren Sorten von Grunddingen. In: *Mathematische Annalen 115*. S. 485-506, 1938.
- [Sch95] Schöning, U.: *Logik für Informatiker*. Spektrum Akademischer Verlag, Heidelberg, 1995.
- [Sch96] Schmitt, A.: *Einführung in graphisch-geometrische Algorithmen*. Teubner, Stuttgart, 1996.
- [Sch99] Schmidtke, H. R.: *Ein formaler Ansatz für die relative Lokalisierung ausgedehnter Objekte in 2-dimensionalen Layouts*. Diplomarbeit am FB Informatik der Universität Hamburg, 2001.
- [ScL83] Schmolze, J.G., Lipkis, T.A.: Classification in the KL-ONE Representation System. In: [Bun83], S.330-332.
- [SiW99] Simon, H.A., Wallach, D.: Cognitive modelling in perspective. In: *Kognitionswissenschaften 9*. S. 1-4, 1999.
- [ScR88] Schnelle, H., Rickheit, G.: *Sprache in Mensch und Computer*. Westdeutscher Verlag, Opladen, 1988.
- [Sha88] Shastri, L.: *Semantic networks: An evidential formalization and its connectionist relation*. Pitman, London, 1988.
- [Sow91] Sowa, J.F. (Hrsg.): *Principles of semantic networks: Explorations in the representation of knowledge*. Morgan Kaufmann Publishers, Inc., San Mateo, 1991.
- [STE00] Schmidtke, H.R., Tschander, L., Eschenbach, C., Habel Ch.: Change of Orientation. In: [ZeS00] S. 57-67.
- [TSH03] Tschander, L.; Schmidtke, H.R.; Habel, Ch.; Eschenbach, C.; Kulik, L.: A geometric agent following route instructions. In: [FBH03] S. 89-111.
- [TvL99] Tversky, B., Lee, P.U.: On pictorial and verbal tools for conveying routes. In: [FrM99], S. 51-64.
- [TWB97] Trullier, O., Wiener, S.I., Berthoz, A., Meyer, J.-A.: Biologically Based Artificial Navigation Systems: Review and Prospects. In: *Progress in Neurobiology, Vol. 51*, S. 483-544, 1997.

- [WKH00] Werner, S., Krieg-Brückner, B., Herrmann, T.: Modelling Navigational Knowledge by Route Graphs. In: [FBH00], S. 295-316.
- [WKM97] Werner, S., Krieg-Brückner, B., Mallot, H.A., Schweizer, K., Freksa, C.: Spatial Cognition: The Role of Landmark, Route, and Survey Knowledge in Human and Robot Navigation. In: [JPP97], S. 41-50.
- [ZeS00] van der Zee, E., Slack, J. (Hrsg.): *Proceedings of the Workshop on Axes and Vectors in Language and Space*. University of Lincolnshire and Humberside, 2000.
- [ZZL02] J. Zhong, H. Zhu, J. Li, Y. Yu: Conceptual Graph Matching for Semantic Search. In: [Pri02].

Erklärung

Hiermit versichere ich, dass ich diese Diplomarbeit selbstständig durchgeführt habe und dass ich keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 1.12.2003