

Seminar „Routeninstruktionen mit Sprache und Skizze“ WS03/04
Aufgabe zum 15.12.03
Matthias Wester-Ebbinghaus

Teilproblem : Inferenzen für die lokalen Referenzsysteme in der CRIL-Repräsentation

Nach der Phase des “syntactic and semantic processing“ und ersten Schritten der Phase des “instruction processing” (Abb. 2, Abschnitte 3.1 – 3.4) liegen die Routeninstruktionen unterteilt in ihre einzelnen Segmente in ihrer CRIL-Repräsentation bzw. in CRIL-Netzen vor. Mithilfe der Wissensbasis (GCS, “geometric concept specification“) werden die Netze verfeinert mittels Umformungsregeln, die auf axiomatischen Charakterisierungen räumlicher Konzepte beruhen. Trotzdem sind die einzelnen Segmente der Routenbeschreibungen weiterhin getrennt und lokal, vor allem in Bezug auf sie impliziten Referenzsysteme der Segmente, die zu dieser Phase selber Knoten des CRIL-Netzes sind. Durch eine Abfolge von Aktionen, die die temporalen Aspekte der Routeninstruktion darstellen, werden die einzelnen Segmente miteinander in Bezug gesetzt. Es gilt:

1. einzelne Knoten der verschiedenen Segmente (des Pfades als auch in des CRIL-Netzes in verfeinerter Form) als identisch zu identifizieren
2. die lokalen Referenzsystemknoten mit Knoten der durch die Aktionssequenz als vorhergehend definierten CRIL-Teilnetze in Relation zu setzen

Nur bei erfolgreicher Inferenz kann eine adäquate Repräsentation der gesamten Route aufgebaut werden. Sonst bleiben Teilnetze isoliert und ohne Bezug, was eine vollständige Identifikation der Route in der Ausführungsphase i.d.R. unmöglich oder zumindest enorm schwierig macht.

Die erste Aufgabe lässt sich unter der Annahme, dass der Endpunkt eines Segments der Startpunkt des folgenden Segmentes ist, recht problemlos lösen.

Die zweite Aufgabe ist komplizierter wie Tabelle 2 verdeutlicht. Die drei Referenzsysteme rsys2, rsys6 lassen sich über die Richtung der zuletzt genannten Pfade induzieren. Sie werden mittels *projective terms* ohne Landmarkenbezug eingeführt, so dass der zuletzt genannte Pfad der einzig sinnvolle Kandidat als Parameter für das Referenzsystem ist. Die beiden Referenzsysteme rsys4 und rsys7 werden mittels *projective terms* mit Landmarkenbezug eingeführt. In beiden Fällen legt die gewollte Interpretation der Terme (*hinter, vor*) nahe, die Landmarke selber als Ankerpunkt für die Induzierung des Referenzsystems zu wählen. Demgegenüber wird rsys5 ebenfalls mittels eines *projective terms* mit Landmarkenbezug eingeführt. In diesem Fall aber bietet sich nicht die entsprechende Landmarke, sondern wiederum der zuletzt genannte Pfad für die Induzierung des Referenzsystems an.

Es gilt also, Strategien zu entwickeln, abhängig von Parametern wie der Form der *projective terms*, der verwendeten Präpositionen, dem Zusammenhang des Terms mit einer Aktion (ein Pfad, der nur als sichtbar beschrieben wird ohne Aktionsbezug bietet sich in den seltensten Fällen als Anker für ein Referenzsystem an) und eventuell weiteren Parametern die nicht spezifizierten Referenzsysteme zu induzieren.

Teilung der Routeninstruktion in primitive Teilabschnitte

Eine Routeninstruktion ist laut Tschander et al. (2003) angemessen, wenn der Instruierte auf ihrer Basis sein Ziel erreichen kann. Dafür ist es erforderlich, dass der Instruierte die Wegbeschreibung auf seine wahrgenommene Umgebung anpasst und die erforderlichen Schritte bis zu seinem Ziel geht.

In unserem Fall, ist unser Instruierter der beschriebene Geometrische Agent. Aber wie kann man erkennen, wann ein Teilabschnitt einer Routenbeschreibung schon abgearbeitet ist und wo der nächste Teil beginnt? Dies ist im Allgemeinen durch einfache sprachliche Ausdrücke in der Beschreibung erkennbar, aber es müssen auch Fälle berücksichtigt werden können, in welchen bestimmte Aktionen nur angedeutet werden oder in denen solche sprachliche Ausdrücke verwendet werden, die von uns verstanden werden können, aber für einen Agenten Probleme bereiten könnten. Würde man aus der reinen sprachlichen Routeninstruktion heraus versuchen ein internes CRIL-Modell zu erzeugen, könnten Lücken oder Ungenauigkeiten im späteren CRIL-Netz entstehen, was zu einem abweichenden Ziel führen könnte.

z.B. Wegbeschreibung ri_3 : "Man gehe geradeaus, bis man rechts von sich ein mehrstöckiges Gebäude (F) sieht...". Bei dieser Beschreibung, die eigentlich eindeutig scheint, besteht das Problem, dass nicht festgelegt ist, bis wohin man gehen soll, d.h. wo soll ich anhalten? Es wurde zwar das Wort "sehen" als Anhaltspunkt gewählt, jedoch wäre dieses nicht so realisierbar. In diesem Satz liegt eher die Anweisung: "Gehe geradeaus, bis Haus F rechts von dir ist.", d.h. "sehen" muss hier wegfallen.

Meine Idee zur Realisierung und Vereinfachung des Geometrischen Agenten wäre daher, die verschiedenen Routeninstruktionen vorweg zu analysieren und in primitive Teilabschnitte zu zerlegen. Also, bevor das interne CRIL-Modell erstellt wird, werden die Kernaussagen der Routeninstruktion und die damit verbundenen Aktionen herausgearbeitet und diese dann zur Weiterverarbeitung verwendet.

Dadurch wäre es einfacher ein entsprechendes CRIL-Netz zu kreieren und den Geometrischen Agenten an seinen Zielpunkt zu dirigieren.

Diese primitiven Teilabschnitte könnten erzeugt werden durch die Indexierung bestimmter Wörter, die als Anhaltspunkte dienen. Es könnte funktionieren, indem man Wörter die für Landmarken und Bewegungsaktionen stehen aus dem Text heraussucht und entsprechend simple Sätze bildet, wie z.B. "man gehe geradeaus bis (Landmarke)...", "man biege rechts ab, an der (Landmarke)...",

... .

Entsprechendes würde auch mit einer Skizze viel hilfreicher sein. Aber dazu habe ich mir noch nicht genug Gedanken machen können. Aber auf jeden Fall könnte man aus einer Skizze heraus eine ähnliche primitive Beschreibung herausarbeiten, die dann dem Geometrischen Agenten zur Zielfindung hilfreich sein könnte.

Mareile Knees
14.12.2003

Stand:

Analyse der temporalen Strukturen in Routeninstruktionen

Bei der Analyse temporaler Strukturen in Routenbeschreibungen wird es darum gehen, die zeitliche Anordnung und Verortung der in Routeninstruktionen beschriebenen Ereignisse darzustellen. Hierbei werden Mechanismen gesucht, die diese zeitliche Anordnung und Verortung herstellen. Bekannte sprachliche Mittel hierfür sind temporale Präpositionalphrasen „bis Montag“ (die jedoch in dem bisher analysierten Material nicht vorkommen), temporale Adverbien wie in „wo **dann** auf der rechten Seite das Haus E liegt“ und Konjunktionen wie in „Hinter der Überführung nach links abbiegen, **und** der Straße folgen“. Darüber hinaus sollen die Strukturen der einzelnen Ereignisse untersucht werden. Es wird zu prüfen sein, inwiefern die linguistische Klassifizierung von Ereignissen in unterschiedliche aspektuelle Klassen wie *States*, *Activities*, *Accomplishments* und *Achievements* für die hier gewählte Fragestellung hilfreich sein können. Während *States* und *Activities* keinen expliziten Endpunkt einführen, besitzen *Accomplishments* und *Achievements* interne Endpunkte.

Nach bisheriger Analyse scheint die Klassifizierung der Ereignisse in die unterschiedlichen aspektuellen Klassen allein nicht ausreichend zu sein, um die End- bzw. Startpunkte aller Ereignisse zu ermitteln bzw. die zeitliche Anordnung der Ereignisse zu gewährleisten. Die meisten vom Agenten auszuführenden Aktionen werden jedoch als *Activities* beschrieben, bei denen der Endpunkt offen gelassen wird. Daher sollen für die zukünftige Analyse Arbeiten von Ladina Tschander zur Segmentierung von Routeninstruktionen herangezogen werden, in denen unterschiedliche Segmentklassen innerhalb der Routenbeschreibung definiert werden (z.B. Sichtungsaussage, Bewegungsaussage usw.).

Ziel der Analyse der temporalen Strukturen ist es, die Endpunkte und ggf. auch Startpunkte der Ereignisse bzw. der vom Agenten auszuführenden Aktionen aufzufindig zu machen, auch wenn diese nicht explizit benannt werden. Der Agent muss ja spätestens bei der Ausführung der Instruktion wissen, welche Aktionen er gleichzeitig (z.B. sehen und geradeaus gehen) bzw. nacheinander ausführen muss (z.B. abbiegen und der Straße folgen) und wie lange diese Aktionen jeweils andauern. Wie das folgende Beispiel zeigt, sind nicht alle Ereignisse, die in den Routeninstruktionen erwähnt werden, auch Aktionen, die der Agent später auszuführen muss. „Du folgst dieser Linksbiegung und gehst die Straße bis zum Ende, wo dann auf der rechten Seite das Haus E liegt.“ Dennoch enthält auch der letzte Nebensatz wichtige Informationen für das Erreichen des Ziels. Daher gilt es bei der temporalen Analyse soviel Informationen über die vom Agenten auszuführenden Aktionen zu erhalten wie möglich, denn nicht jede Aktion wird auch explizit als Aktion benannt. Die Ergebnisse der temporalen Analyse sollen also dazu genutzt werden, die Segmentierung des Aktionsplans vom GeoAgenten zu unterstützen.

Fragen, die bei der Analyse zu klären sind und die in der letzten Sitzung diskutiert wurden:

- Was wird eigentlich vom Instrukteur an Wissen vorausgesetzt? Über wie viel Weltwissen verfügt der GeoAgent? Und wie viel Wissen über das Gelände wird bei der Analyse der Routeninstruktionen vorausgesetzt?
- Inwiefern stellen eine Ontologie der Ereignisse und der Objekte das Weltwissen des GeoAgenten dar? (→ *Navigation*)
- Um die Segmente bzw. Aktionsarten auch formal zu definieren wird lexikalisches Wissen benötigt (Verb-Info & Argument-Info), d.h. es wird auf das Lexikon und die darin enthaltenen semantischen Primitive zu gegriffen werden (→ *Textuelle Repräsentationen; Representation of sentence meaning*)
- Wenn die Endpunkte bzw. die Startpunkte der einzelnen Aktionen (Ereignisse) bekannt sind, können Teilabschnitte miteinander verknüpfen werden. In der Regel

werden die End- bzw. Startpunkte der Ereignisse (und Aktionen) jedoch nicht explizit angegeben. Daher müssen Mechanismen zu Generierung semantischer Inferenzen definiert werden (\rightarrow *Instruction processing*)

Lösungsansätze aus der Projekt-Literatur:

- Segmentierungsansatz in Fraczak, Lidia (1998). Generating 'mental maps' from route descriptions. In Patrick Olivier & Klaus-Peter Gapp (eds.), *Representation and Processing of Spatial Expressions* (pp. 185–200). Mahwah, NJ: Lawrence Erlbaum.

„Interne Kodierung der Info aus Skizzen“

Stephan Grochowina

Bei der internen Kodierung der Informationen aus Skizzen wird eine Repräsentation aufgebaut, die den Inhalt der Skizzen abbildet und ihre Elemente in Relation zueinander setzt. Auf Grund dieser Repräsentation muss es dann möglich sein, sowohl eine Repräsentation der Welt zu erzeugen, die durch die Skizze beschrieben wird, als auch die Aktionen abzuleiten, die notwendig sind, um von dem zu ermittelnden Startpunkt zum zu ermittelnden Zielpunkt zu gelangen.

Durch die Repräsentation des Skizzeninhaltes kann die Lösung folgender Probleme unterstützt werden:

- Die Skizzenelemente und ihre Anordnung liefern Informationen, um Elemente der Routenrepräsentation (z.B. Landmarken, Wege, Labels, Kreuzungen) zu ermitteln.
- Sie ermöglichen die Zuordnung von Beschriftung.
- Sie unterstützen die 3D-Darstellung.

Im ersten Schritt werden die Grafikprimitive der Skizze eingelesen (z.B. Pixel (x,y) , die eventuell noch mit dem Zeitpunkt t ihrer Erzeugung verknüpft sind). Aus diesen werden dann höhere Objekte unter Verwendung der räumlichen und eventuell zeitlichen Relationen zusammengesetzt (z.B. werden alle räumlich, bzw. zeitlich „nahen“ Pixel zu einem Objekt zusammengefasst), die möglicherweise wieder zu Objekten zusammengefasst werden. Diese Objekte weisen gewisse Eigenschaften auf (wo in der Skizze befinden sie sich, wie beschaffen sind sie, etc.) und stehen in räumlicher (bzw. auch zeitlicher) Relation zueinander. Dadurch ergeben sich folgende Fragen:

- Welche Objektarten/Skizzenelemente gibt es? Welche Eigenschaften haben sie? Wie werden sie aus niederen Objekten/Grafikprimitiven gebildet?
- In welchen Relationen können sie zueinander stehen?
- Wie kann dies repräsentiert werden?

Problemdarstellung: Start- und Endpunkte von Wegen in Skizzen

Nils Andresen

NilsAndresen@web.de

16. Dezember 2003

Darstellung

- Wie kodiert man
 1. Eindeutig
 2. Intuitivden Start- bzw. Endpunkt einer Skizzen-Wegbeschreibung (Depiktion)
- Wie kodiert man „Den Weg“ einer Depiktion

Weiter aufkommende Fragen

- ist jeder Start- auch Endpunkt ?
- auch wenn ein Weg aus Teilwegen bestehen kann (Follow-Kown-Route) ?
- gibt es „Unteilbare“ Wegstücke ?
- Muß „Der Weg“ in einer Depiktion explizit kodiert sein ?
 - Mögliche Alternativen von Wegen
 - Erkennung aller möglichen Wege und Berechnung des „Besten“

Vorausgegangene Lösungsansätze

- Kreuze (o.ä.) mit dem Label „Start“ bzw. „Ziel“
- Kodierung von Wegen durch Striche oder Doppelstriche bzw. Farblich (Stadtplan)
- Kodierung „Des Weges“ durch Pfeile

Schritte zur Lösung¹

Folgende Teilschritte bieten sich an:

- Erkennen von Speziellen (evtl. gefärbten o.ä.) Karten,
- Erkennen von „echten“ Skizzen.

¹In diesem Papier beachte ich nicht die in der Sitzung besprochene Alternative des Toolkits.

Seminar „Routeninstruktionen mit Sprache und Skizze“ WS03/04
Aufgabe zum 15.12.03

Martin Malzahn

Das Problem: Wann gehört eine Beschreibung zu einem Objekt?

Welche Beschreibung gehört zu welchem Objekt?

Und die wichtige Frage die sich dann stellt: Was ist ein Objekt überhaupt?

Bei jeder Skizze, in der auch Beschreibungen verwendet werden, müssen diese einzelnen Objekten zugeordnet werden. Doch nach welchen Kriterien geschieht dies? Mal steht eine Beschreibung in einem Objekt, mal an einem, dann wieder an einer Gruppe oder Menge von Objekten und gelegentlich an einer Region. Wie ordnet nun also ein Mensch diese Beschreibungen den passenden Objekten zu, falls es denn welche gibt? Spielen dabei Kriterien wie räumliche Nähe eine Rolle? Und wie könnte ein Agent das bewerkstelligen?

Dieses Problem tritt dann in ähnlicher Form auch bei natürlich-sprachlichen Beschreibungen von Routen auf. Wann ist ein Haus groß? Wie kann ein Mensch und später ein Agent entscheiden, ob ein Attribut beziehungsweise eine Beschreibung auf ein Objekt zutrifft, die nicht sonderlich genau ist und eventuell mehrere Objekte in Betracht kommen lässt? Was ist überhaupt ein Objekt? Ist eine Straße ein Objekt? Oder ein Haus? Was ist ein Platz? Wie sind Objekte begrenzt und woran erkennt man diese Grenzen?

Dann ist da noch das Problem, dass auf einigen Ebenen einzelne Dinge Objekte sind oder zumindest deren Stellenwert einnehmen auf anderen Ebenen diese Dinge diesen Status dann aber verlieren. Beispiel: Drei Bäume könnten auf einer Beobachtungsebene drei unterschiedliche Objekte sein auf einer anderen Ebene ist es nicht wichtig, dass es drei Bäume sind, sondern es zählt nur noch, dass es ein Objekt 'Baumgruppe' gibt, auf der nächsten Ebene sind diese Bäume gar nicht mehr wichtig und werden nicht mehr als Objekte oder Objekt deklariert.

Dann wiederum gibt es Situationen in denen ein Objekt in mehrere unterschiedliche Objekte gespalten wird oder mehrere Objekte zu einem zusammengefasst werden, weil man neues Wissen hinzugewonnen hat und somit einige als gegeben angenommene Vorstellungen ihre Gültigkeit verlieren. Man sieht also: bei einem Perspektivenwechsel oder Veränderung des Wissens kann sich der Status eines Objektes ändern.

Kann der Agent Beschreibungen nicht den richtigen Objekten zuordnen oder findet er erstmal gar keine Objekte, denen er überhaupt Beschreibungen zuordnen kann, helfen diese ihm nicht bei der Wegfindung oder führen ihn gar auf einen falschen Weg, falls einige der Beschreibungen falsch zugeordnet werden.

Um Objekte und Beschreibungen einander zuordnen zu können, muss der Agent natürlich erstmal erkennen, was denn überhaupt Objekte sind, sowohl in der Skizze und der natürlich-sprachlichen Beschreibung als auch in der Umwelt, in der er sich dann später bewegt. Er muss also erstmal mögliche Objekte aus den Beschreibungen

'heraus filtern', um diese dann auf später in seiner Umwelt erkannten Objekte 'abzubilden'.

Wie könnte man dieses Problem nun also angehen? Zuerst einmal müsste man eine grobe Idee davon bekommen wie Menschen diese Probleme lösen, falls sie sie denn lösen. Dann müsste man ein geeignetes Modell finden, um dem Agenten diese Fähigkeit zugänglich zu machen.