

Hallesches Jahrb. Geowiss.	R. A	Bd. 26	Halle (Saale) 2004	S.
----------------------------	------	--------	--------------------	----

Ausgewählte Methoden der Künstlichen Intelligenz in der Geographie

Selected Artificial Intelligence techniques in Geography

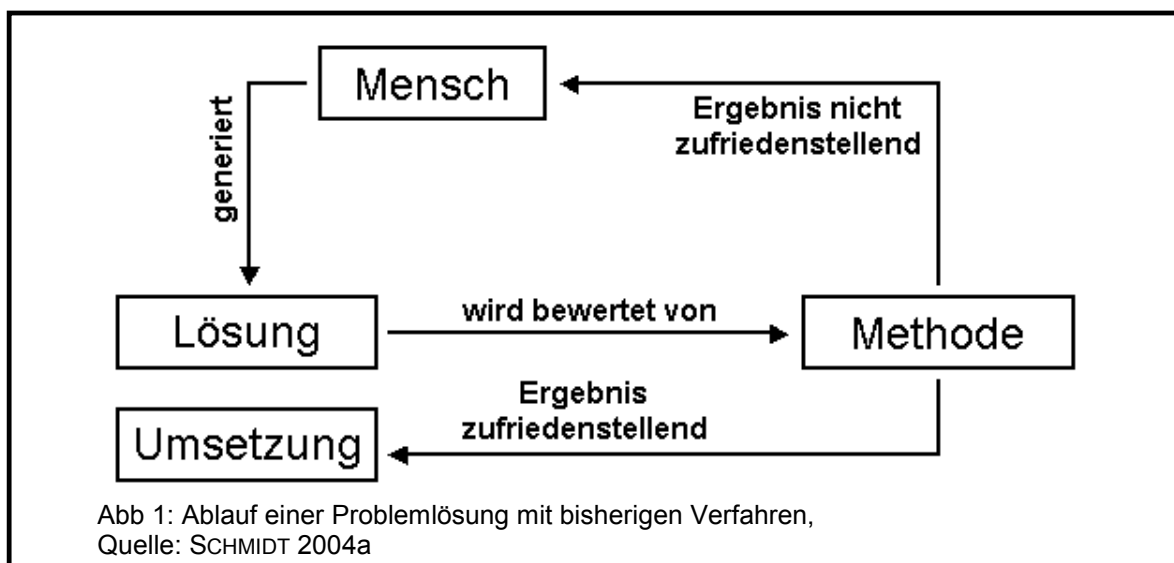
Mit 5 Abbildungen

VON THORSTEN SCHMIDT UND HANS-JOACHIM ROSNER

Zusammenfassung: Die bisherige Arbeitsweise der Geographie zur Lösung von Praxisproblemen orientiert sich in der Regel an der Überprüfung von Lösungsalternativen. Zu einer Problemstellung werden Lösungsvarianten entworfen und auf ihre Leistungsfähigkeit hin untersucht. Die Qualität der gefundenen Lösungen hängt vom Fachwissen des jeweiligen Bearbeiters ab und unterliegt den Beschränkungen menschlicher Vorstellungskraft und Beurteilungsgabe. Das Forschungsgebiet der Künstlichen Intelligenz stellt ein reichhaltiges Methodenset zur Verfügung, welches die Problemlösungsfähigkeit der Geographie erheblich steigern kann. Durch Einsatz dieser, in der Geographie bisher kaum verbreiteten Techniken, werden Computer in die Lage versetzt, selbst optimale oder zumindest sehr gute Lösungen zu Praxisproblemen zu finden.

Summary: The common methods of geography concerning the solution of practical problems usually refer to a checklist of alternative solutions. Variant solutions are listed and examined according to their applicability. The quality of a solution depends on the expertise of the person in charge and is subject to the limited human power of imagination and judgement. The research domain of Artificial Intelligence provides a rich methodological set of techniques which may increase the efficiency of problem solving in geography. By employing such techniques to the domains of geography - a step that has scarcely been made up to the present - computers are enabled to find optimal or at least very close to optimal solutions to pragmatic problems independently.

1 Einführung



Insbesondere die Angewandte Geographie beschäftigt sich mit der querschnittsorientierten Anwendung geographischen Wissens auf Probleme der Gesellschaft, Politik und Wirtschaft. Sie stellt Methoden zur Lösung von Praxisproblemen mit räumlicher Relevanz zur Verfügung. Neben qualitativen

Methoden kommen seit Ende der 70er Jahre stochastische Verfahren und Bewertungsmethoden zum Einsatz. Auch die Verwendung Geographischer Informationssysteme zur Problemanalyse und Problemlösung hat sich etabliert und stellt einen weiteren methodischen Baustein dar. Allen diesen quantitativen Methoden ist die Tatsache gemein, daß sie statisch und strukturell festgelegt sind. Zu einem Problem aus der Praxis werden mit Hilfe menschlichen Sachverständs, Erfahrungswerten und Annahmen Umsetzungs- oder Lösungsmöglichkeiten entworfen. Die so gefundenen Varianten werden daraufhin mit den vorhandenen Methoden bewertet und gegebenenfalls umgesetzt (Abb. 1). Diese Vorgehensweise ist unbefriedigend und entspricht nicht dem Stand der technischen Möglichkeiten.

Zwei Beispiele sollen die Problematik verdeutlichen: Ein GIS ist durchaus in der Lage, die Schallausbreitung einer geplanten Straße zu modellieren und den Lärmpegel in einem benachbarten Wohngebiet zu ermitteln. Es ist aber nicht in der Lage, eine Trasse zu finden, bei dem der störende Lärm in Wohngebieten minimiert wird. In der gängigen Praxis verbleibt diese Aufgabe beim menschlichen Sachverständ. Ein weiteres Beispiel: Der Gesetzgeber schreibt vor, daß die Rettungsdienste jeden Ort des Landes innerhalb von 15 Minuten nach Eingang eines Notrufs erreichen müssen. Aus Kostengründen wird versucht, die Anzahl der Rettungswachen so gering wie möglich zu halten. Bisher gibt es keine Methode, die dieses Problem zufriedenstellend lösen kann. Vielmehr müssen die Standorte der Wachen nach Gutdünken gesetzt werden. Anschließend läßt sich ermitteln, ob die Vorgabe, jeden Ort innerhalb von 15 Minuten zu erreichen, erfüllt wurde. Offensichtlich handelt es sich um ein Optimierungsproblem, welches mit verfügbaren Methoden nicht befriedigend zu lösen ist. Im Folgenden soll gezeigt werden, daß es Methoden aus dem Bereich der Künstlichen Intelligenz gibt, die zumindest theoretisch das Potential haben, diese und andere Probleme besser als bisher zu lösen.

2 Künstliche Intelligenz

Künstliche Intelligenz (KI), als Teilgebiet der Informatik versucht Maschinen zu entwickeln, die sich so verhalten, also ob sie über Intelligenz verfügen würden. Es geht also nicht zwingend darum, menschliche Denkweisen in eine Maschine zu übertragen, sondern mit Hilfe bestimmter Verfahren Intelligenz zu simulieren. Das Forschungsgebiet der Künstlichen Intelligenz verfügt über ein ausgedehntes Methodenset und stellt vielen Disziplinen nützliche Verfahren zur Verfügung. Expertensysteme unterstützen medizinische Diagnosen, Mustererkennung spielt in der Robotik eine große Rolle und die Computerspiel Industrie setzt auf KI-Techniken, um den Menschen mit "intelligenten" Spielgegnern herauszufordern. Die Methoden der KI-Forschung sind umfangreich: Suchverfahren, Fuzzy-Logik, Neurocomputing, Expertensysteme, genetische Verfahren oder Artificial Life sind nur einige Beispiele. Die deutschsprachige Geographie hat sich mit den Möglichkeiten, die sich durch die rasche Entwicklung leistungsfähiger Computer ergeben, bisher nur spärlich befaßt. Dies spiegelt sich in der Literaturlage zur Thematik wider. Deutschsprachige Publikationen, die sich ausdrücklich mit Künstlicher Intelligenz in der Geographie beschäftigen, gibt es kaum. Nur vereinzelt finden sich geographische Publikationen, die sich mit Fuzzy Logik, neuronalen Netzen oder agentenbasierten Systemen beschäftigen. Für den angelsächsischen Sprachraum hat OPENSHAW 1997 mit dem Titel "Artificial Intelligence in Geography" ein grundlegendes Werk veröffentlicht.

Über die Frage, warum die deutschsprachige Geographie sich bisher sehr wenig mit diesem Themengebiet beschäftigt hat, läßt sich nur spekulieren. Eventuell hat die GIS-Revolution seit Mitte der 80er Jahre alle Aufmerksamkeit auf sich gezogen. Trotzdem ist es überraschend, daß Methoden, die sich offensichtlich in anderen Fachgebieten bewährt haben, von Geographen bisher unberücksichtigt geblieben sind. Auch dieser Aufsatz kann keine funktionsfähigen Anwendungen für die Geographie präsentieren. Aber er will zeigen, daß KI über große Potentiale zur Problemlösung verfügt. Die größte Schwierigkeit für einen Geographen, der sich mit KI-Methoden beschäftigt, mag darin bestehen, sich mit Literatur zu befassen, die von Computerwissenschaftlern mit ihrem eigenen, teilweise sehr speziellen und mathematischem Vokabular geschrieben wurde.

3 KI-Methoden

Bei den hier vorgestellten Verfahren handelt es sich ausschließlich um Suchverfahren. Viele Probleme insbesondere der Angewandten Geographie können als Suchprobleme formuliert werden. Erst das Verständnis von Suchalgorithmen ermöglicht die Aussage, ob ein Problem aus der Praxis mit ihrer Hilfe lösbar ist. Bevor ein Suchverfahren angewendet werden kann, muß das Problem in eine Form gebracht werden, die eine Verarbeitung durch ein Suchverfahren ermöglicht. Oftmals geschieht dies durch die Repräsentation des Suchproblems in Gestalt eines Suchbaums. Der nächste Schritt besteht in der computergerechten Definition einer guten Lösung des Problems. Sie ist erforderlich, um die Lösungsvorschläge, die der Computer erarbeitet, zu bewerten und gegebenenfalls den Suchprozeß in

die richtige Richtung zu lenken. Nur wenn diese Bedingungen erfüllt sind, kann einer der nachfolgenden Methoden angewandt werden. Welches das geeignetste Verfahren ist, hängt von der jeweiligen Problemstellung ab und muß im Einzelfall entschieden werden. Die hier erläuterten Methoden erheben keinen Anspruch auf Vollständigkeit. Sie sollen lediglich das Anwendungspotential der KI-Methoden illustrieren. Die Anwendungsbeispiele sind abstrakt und theoretisch gehalten, um den Blick auf das wesentliche – die Methode – nicht zu verstellen.

Die hier beschriebenen Beispiele wurden mit der Software NetLogo (WILENSKY 1999) realisiert. Net-Logo ist ein agentenbasiertes Programm, welches ursprünglich zur Simulation sozialer Phänomene und Prozesse entwickelt wurde. Lauffähige Java-Applets dieser und anderer Suchmethoden finden sich bei SCHMIDT 2004b.

3.1 Blinde Suche

Die Blinde Suche ist ein Verfahren, welches auf den ersten Blick am wenigsten intelligent erscheint. Um ein Problem zu lösen wird massive Rechenkapazität eingesetzt. Das Computerprogramm generiert alle theoretisch möglichen Lösungen zu einem Suchproblem und bewertet sie mit Hilfe einer eingebauten Evaluierungsfunktion. Sie stellt eine mathematische Formulierung der Ansprüche an eine gute Lösung dar. Diejenige Lösung, für die die Evaluierungsfunktion den größten Wert ermittelt hat, wird ausgegeben. Das Vorgehen erinnert an das Knacken eines Zahlenschlosses. Je mehr Rechenkraft und –zeit zur Verfügung steht, desto komplexere Schlösser können geöffnet werden. Übertragen auf eine geographische Suche bedeutet dies, daß Probleme, die eine endliche und überschaubare Anzahl von theoretischen Lösungsmöglichkeiten besitzen auf diese Weise optimal gelöst werden können. Für kombinatorische Probleme ist das Verfahren ungeeignet, da die Zahl der theoretischen Möglichkeiten exponentiell oder gar faktoriell ansteigt. Der wesentliche Vorteil der Methode besteht darin, daß sie verlässlich das optimale Ergebnis im Sinne der Evaluierungsfunktion ermittelt.

Anwendungsmöglichkeit:

Gedacht sei ein besiedelter Raum, der möglichst effektiv mit Bushaltestellen erschlossen werden soll. Als Datengrundlage wird eine gerasterte Bevölkerungsdichtekarte verwendet. Jeder Bildpunkt repräsentiert eine bestimmte Anzahl von Einwohnern. Es wird angenommen, daß Haltestellen nur bis zu einer maximalen Luftlinienentfernung X vom Wohnort zur Haltestelle akzeptiert werden. Gesucht werden die optimalen Standorte für die Haltestellen. Die Zielvorgabe lautet mit jeder Haltestelle so viele Personen wie möglich zu erreichen. Im ersten Schritt wird sequentiell für alle Bildpunkte die Anzahl der Personen berechnet, die im Radius X um ihn herum leben. Ist diese Berechnung für alle Rasterzellen abgeschlossen, wird die erste Haltestelle an dem Punkt gesetzt, an dem dieser Wert am höchsten ist. Daraufhin wird die Bevölkerung im Radius X um die erste Haltestelle herum auf Null gesetzt. Diese Personen sind versorgt und dürfen bei der Berechnung der zweiten Haltestelle nicht mehr in die Suche mit einbezogen werden. Für die zweite und alle weiteren Haltestellen wiederholt sich dieser Prozeß. Unter den gestellten Bedingungen ist das erzielte Ergebnis optimal in dem Sinne, daß keine Verbesserungen möglich sind. In der Wirklichkeit können Bushaltestellen aber nicht beliebig gesetzt werden. Diese Tatsache würde die Suche einfacher und schneller machen, da dann nur noch die Rasterzellen, die an einer Straße liegen, untersucht werden müßten. Vielleicht wäre eine Haltestelle an einem Bahnhof nach dem beschriebenen Suchverfahren gar nicht optimal, jedoch ist sie zwingend erforderlich um die Anbindung von Bus und Bahn zu ermöglichen. Innerhalb des Programms ist es ein leichtes, bestimmte Haltestellen unverrückbar festzulegen. Das Problem der optimalen Route für eine Buslinie ist damit aber noch nicht gelöst. Ein genetisches Verfahren wie unter Punkt 3.4 beschrieben ist hierfür ein brauchbarer Ansatz.

3.2 Heuristiken

Viele Suchprobleme sind derart komplex, daß sie sich mit einer blinden Suche nicht in praktikabler Zeit lösen lassen. Grund hierfür ist die exponentiell oder faktoriell ansteigende Anzahl der zu untersuchenden Möglichkeiten. Beispielhaft soll dies an der Suche nach der besten Standortverteilung für 100 Mobilfunkmasten auf 1000 mögliche Standorte illustriert werden. In diesem Fall gibt es

$$\frac{1000!}{100!}$$

Möglichkeiten – eine Zahl, die sich aufgrund ihrer Größe mit gängiger Software weder berechnen noch darstellen läßt.

Heuristische Suchverfahren versuchen der Problematik dadurch zu entgehen, daß sie auf eine vollständige Untersuchung des Suchraumes verzichten. Dies kann allerdings zur Folge haben, daß entweder gar keine Lösung gefunden wird oder die gefundene Lösung eventuell nicht optimal ist. Heuristiken können nach folgenden Kriterien unterschieden werden:

- erschöpfend – nicht erschöpfend
- informiert – uninformiert

Ein erschöpfendes Verfahren findet garantiert die optimal Lösung, obwohl es nicht zwingend – wie bei der blinden Suche - den vollständigen Suchraum untersucht. Als Beispiel soll eine Namensdatenbank dienen, die zu jedem Eintrag Name, Vorname und Geburtsdatum enthält. Sucht ein Anwender in der Datenbank nach dem Geburtsdatum von Peter Müller, so wird er erwarten, alle Einträge angezeigt zu bekommen, die in der Spalte "Name" den Begriff "Müller" und in der Spalte "Vorname" den Begriff "Peter" enthalten. Die Software, die die Datenbank verwaltet, könnte nun alle Datensätze auf diese Kriterien hin durchsuchen. Allerdings wäre es effektiver, nur bei denjenigen Datensätzen die Spalte "Vorname" zu untersuchen, in denen die Überprüfung der Spalte "Name" bereits den Begriff "Müller" ergeben hat. Obwohl nicht alle Datensätze beziehungsweise Suchpfade vollständig untersucht wurden, wird das vollständige und korrekte Ergebnis ausgegeben werden. Die Suchzeit ist gegenüber der vollständigen Suche in etwa halbiert. Übertragen auf das Suchen in Suchbäumen bedeutet dies, daß Pfade, die offensichtlich abwegig sind, da sie zum Beispiel eine Bedingung verletzen oder die Suchkosten bereits einen bestimmten Wert überschritten haben, von der weiteren Suche ausgeschlossen werden.

Ein nicht erschöpfendes Verfahren kann ein Ergebnis liefern, ohne daß der Anwender die Gewißheit hat, daß es nicht doch eine bessere Lösung gibt. So wird zum Beispiel ein Geschäft gesucht, welches ein Produkt zum günstigsten Preis verkauft. Der Kaufwillige erkundigt sich daraufhin in fünf Geschäften der Stadt und kauft das Produkt in dem Laden, in dem er es am günstigsten bekommt. In der Regel hat der Käufer nun ein gutes Gefühl, obwohl er nicht sicher sein kann, ob er das Produkt in einem der fünf anderen Läden der Stadt (oder gar im Nachbarort) nicht hätte billiger bekommen können.

Bei informierten Verfahren liegt für jeden Suchschritt ein heuristischer Schätzwert vor, der auf seine Güte, bzw. seine Nähe zum Lösungszustand schließen läßt. Ein Bergsteiger, der den Weg zum Gipfel sucht, wird den nächsten Schritt in der Regel bergauf setzen, da er weiß, daß er zum Erreichen seines Ziels bergauf gehen muß. Sollte er nach einiger Zeit feststellen, daß der von ihm eingeschlagene Weg eine Sackgasse ist, wird er umkehren und einen anderen Weg wählen. Eben so verhält es sich bei den informierten Verfahren. Der vielversprechendste Pfad wird zuerst untersucht. Führt er nicht zum Ziel, kommt der zweitbeste Pfad an die Reihe.

Nicht informierte Verfahren verfügen über keine heuristische Schätzfunktion. Bis zum Erreichen der Lösung haben sie keinen Anhaltspunkt dafür, wie gut der eingeschlagene Pfad ist. Beispielhaft hierfür ist die Suche nach der Nadel im Heuhaufen. Der Suchende weiß nicht, ob es besser ist, links oder rechts, oben oder unten zu suchen. Erst wenn er die Nadel gefunden hat, ist er sich über den besten Lösungsweg im Klaren.

Heuristische Suchverfahren ermöglichen also Lösungsverfahren für Probleme, die unter Verwendung der Blinden Suche erst mit der nächsten oder übernächsten Rechnergeneration lösbar wären. Die verfügbare Anzahl heuristischer Suchverfahren ist wesentlich größer, als die hier vorgestellten (siehe zum Beispiel WINSTON 1992). Je nach Problemstellung, Anforderung an die Güte der Lösung oder die verfügbare Rechenzeit können unterschiedliche Verfahren eingesetzt werden.

3.3 Monte-Carlo-Suche

Die Idee hinter der Monte-Carlo Methode ist einfach. Für ein Suchproblem werden tausende bis abertausende zufällige Lösungsvorschläge generiert. Für jeden Vorschlag wird die Leistungsfähigkeit anhand einer Evaluierungsfunktion ermittelt und die jeweils beste gefundene Lösung gespeichert. OPENSHAW (OPENSHAW & OPENSHAW 1997 : 102) meint hierzu: „*It is crude, but can be surprisingly effective.*“

Das gegebene Problem muß auf irgendeine Art und Weise parametrisierbar sein. Der Lösungsgenerator ist so zu programmieren, daß die zufällig erstellten Lösungen das Problem zumindest theoretisch auch zufriedenstellend lösen können. Die Qualität der erzielten Lösung hängt von der investierten Rechendauer ab. Ein wesentlicher Vorteil der Methode ist, daß sie parallelisiert werden kann, das

heißt, daß auf mehreren Rechnern gleichzeitig nach Lösungen gesucht werden kann. Da das Verfahren sehr einfach umzusetzen ist, mag ein Versuch in jedem Fall lohnen. Dabei ist zu berücksichtigen, daß die Methode weder informiert noch erschöpfend ist. Der Nachweis, daß es keine bessere Lösung gibt, als die bisher gefundene, kann nicht erbracht werden. Die Anwendungsmöglichkeiten sind vielfältig. Sie reichen von der Parameteroptimierung bis zur Standortsuche.

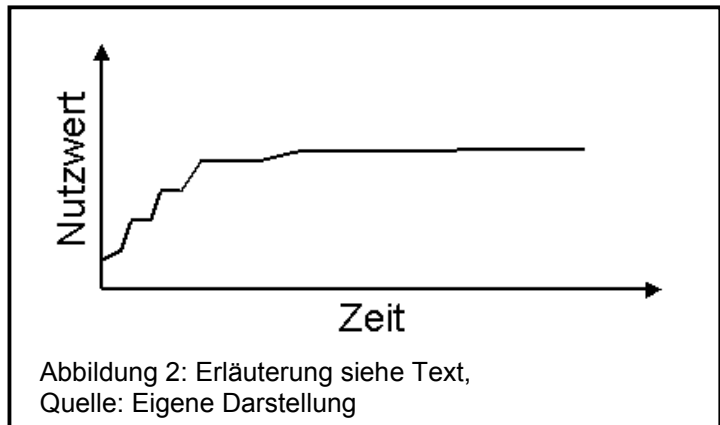
Anwendungsmöglichkeit:

Die Positionierung von Mobilfunkmasten ist in der Realität ein sehr komplexer und schwieriger Prozeß. Neben technischen Aspekten und gesetzlichen Regelungen gilt es, zwischen Ängsten der Bevölkerung und optimaler Signalabdeckung Lösungen zu finden, die mehreren Interessen gerecht werden. Das folgende stark vereinfachte Beispiel vernachlässigt alle diese Aspekte, da es vielmehr um die Problematik geht, mehrere Einrichtungen, deren Standorte sich gegenseitig beeinflussen, optimal im Raum zu verteilen. Statt Mobilfunkmasten hätten auch Einzugsbereiche von Einkaufszentren oder Zuständigkeitsbereiche von Polizeidienststellen als Beispiel dienen können.

In einem gegebenen Raum sollen eine bestimmte Anzahl von Mobilfunkmasten so verteilt werden, daß eine möglichst hohe Abdeckung des Gebiets erreicht wird. Wie bereits erwähnt, gibt es für die Verteilung von X Masten auf einer Anzahl von Y möglichen Standorten

$$\frac{Y!}{X!}$$

unterschiedliche Möglichkeiten. Aus diesem Grund ist eine Blinde Suche ausgeschlossen, da sie in menschlichen Zeiträumen nicht abzuschließen wäre. Ein Lösungsgenerator positioniert die Masten im ersten Schritt zufällig über die Karte. Sind alle Masten gesetzt, wird die Signalausbreitung modelliert und die Güte der Positionierung über eine Evaluierungsfunktion nutzwertanalytisch ermittelt. Dazu wird der Signalabdeckung in den unterschiedlichen Landnutzungen ein Gewicht zugewiesen. Es liegt nahe, der Signalabdeckung in Wohngebieten und auf Straßen ein höheres Gewicht zuzuweisen, als anderen Flächen. Der so ermittelte Nutzwert beschreibt die Güte der Lösung. Dieser Prozeß wird nun beliebig oft wiederholt. Diejenige Lösung, die den höchsten Nutzwert erzielt hat, wird gespeichert und während des Programmablaufs gegebenenfalls durch neuere, bessere Lösungen ersetzt. Die größten Verbesserungen des Nutzwertes werden zu Beginn erzielt. Da die Wahrscheinlichkeit mit jeder gefundenen besseren Lösung sinkt, eine noch bessere Variante zu finden, verlängern sich die Intervalle zwischen den Qualitätssprüngen. Abb. 2 zeigt einen typischen Entwicklungsverlauf der Qualität bisher gefundener bester Lösungen. Um die Ergebnisse zu verbessern, wäre eine Weiterentwicklung der besten Lösung mit Hilfe eines genetischen Verfahrens sinnvoll und möglich (siehe Kapitel 3.5).



3.4 Kostenbasierte Suche

Eine Möglichkeit, die Suche innerhalb eines großen Suchraums in eine plausible Richtung zu lenken ist der A* Algorithmus. Er dient zur Ermittlung des kostengünstigsten Pfades zwischen einem Start- und einem Zielknoten innerhalb eines Suchbaums. Das Verfahren bezieht sich ursprünglich auf die Suche innerhalb von Suchbäumen, aber auch Rasterkarten können als eine alternative Schreibweise für einen Suchbaum verwendet werden. Jeder Bildpunkt stellt einen Knoten dar, der mit allen benachbarten Bildpunkten verbunden ist. Es ist ein erschöpfendes Verfahren, welches nur auf Suchprobleme anwendbar ist, bei denen bereits während der Suche plausible Annahmen darüber getroffen werden können, wie weit man vom Erreichen des Lösungszustandes noch entfernt ist. Beim Expandieren der Knoten innerhalb eines Suchbaumes werden die Kosten aller neuen Knoten mittels der Funktion

$$K = K_b + H$$

berechnet. K steht für die zu berechnenden Kosten, K_b für die Summe der Kosten, die bisher aufgelaufen sind, um den jetzigen Knoten zu erreichen und H für eine heuristische Schätzfunktion. Sie gibt den Wert der minimalen Kosten vom aktuellen Knoten zum Zielknoten an. Wichtig hierbei ist, daß die

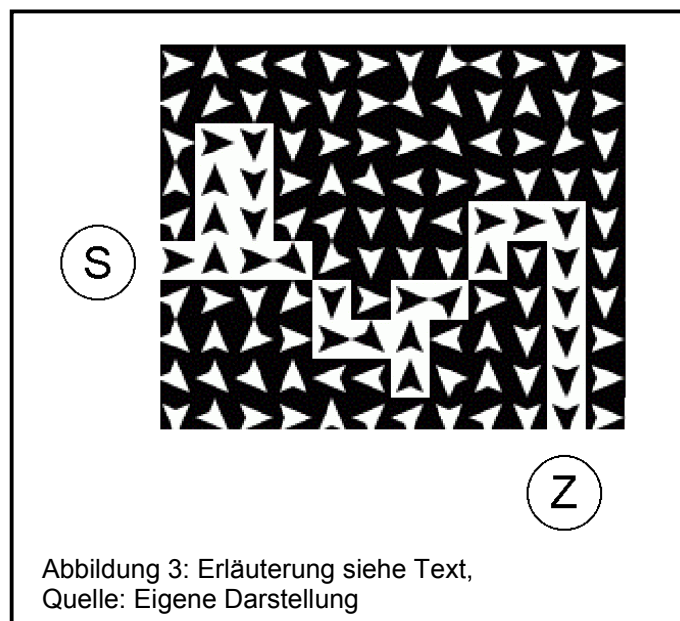
Schätzfunktion die untere Schranke bildet, also wirklich die theoretisch minimalen Kosten zur Erreichung des Zielknotens verwendet, da die Suche sonst in die Irre geleitet werden kann.

Der eigentliche Ablauf des Algorithmus ist einfach: Bei jedem Schritt wird derjenige Knoten expandiert, dessen Kosten K am geringsten sind. Soll neben den Kosten des günstigsten Pfades auch der Pfad selbst gespeichert werden, so ist neben den Kosten auch der Vorgänger eines jeden Knotens zu speichern.

Anwendungsmöglichkeit

Das kostenbasierte Vorgehen des A* Algorithmus erlaubt die Kombination mehrerer Kostenarten. Diese Eigenschaft soll im folgenden Beispiel dazu dienen, eine optimale Trasse für eine Infrastruktureinrichtung (zum Beispiel eine Straße) zwischen zwei Punkten in einem vorgegebenen Raum zu finden. Als Datengrundlage dient eine Landnutzungskarte, der ein Höhenmodell unterlegt ist. Zur Trassensuche werden drei verschiedene Arten von Kosten herangezogen. Die "Distanzkosten" beziehen sich auf die Länge der Trasse. Die Kosten zur Überwindung unterschiedlicher Landnutzungen werden vereinfacht als "Ökokosten" bezeichnet. Sie entsprechen in etwa den ökologischen Wertigkeit einer Fläche. So ist eine Trassenführung über eine Ackerfläche sehr kostengünstig wohingegen das Passieren eines Landschaftsschutzgebiets zwar theoretisch möglich ist, aber nur zu vergleichsweise sehr hohen Kosten. Der letzte Kostenfaktor bezieht sich auf die Überwindung von Höhendifferenzen. Große Steigungen oder Gefälle verursachen hohe Kosten; im ebenen Gelände wird dieser Kostenfaktor Null. Da keine allgemein gültige Theorie oder Methode bekannt ist, diese unterschiedlichen dimensionslosen Kosten miteinander zu verrechnen, muß der Anwender über die Gewichtungen der einzelnen Kostenfaktoren selbst entscheiden.

Gesucht ist nun der kostengünstigste Weg von einem Startpunkt S zu einem Zielpunkt Z. (Abb. 3). Dazu expandiert der Startpunkt Knoten in alle möglichen Richtungen. Die Kosten für jeden Knoten werden nach obiger Formel berechnet. Als heuristische Schätzfunktion H wird die euklidische Distanz zum Zielpunkt verwendet. Sie bildet garantiert eine untere Schranke, da kein günstigerer Pfad möglich ist. Im nächsten Schritt wird der kostengünstigste Knoten wiederum in alle Himmelsrichtungen expandiert; alle Knoten werden bewertet und der günstigste expandiert weiter. Nach jeder Expansion richten sich die Knoten auf ihren kostengünstigsten Nachbarn aus. Dieser Prozeß wird so lange wiederholt, bis der Zielknoten erreicht ist. Abb. 3 zeigt das Ergebnis eines Suchvorgangs. Die Landnutzung und das Höhenmodell konnten aus drucktechnischen Gründen nicht wiedergegeben werden. Die Dreiecke repräsentieren die Knoten. Jeder Knoten zeigt auf seinen kostengünstigsten Nachfolgeknoten. Die optimale Route ist weiß unterlegt.



Wie bereits angesprochen, sind etliche methodische Schwierigkeiten nicht gelöst. Problematisch ist nicht nur die Verrechnung und Gewichtung der unterschiedlichen Kostenarten, sondern auch die Ermittlung der Kostenhöhe selbst. Die Entwicklung einer solchen, fachlich begründeten Methode bleibt eine eigene wissenschaftliche Arbeit. Die technischen Möglichkeiten einer Umsetzung hierfür sind gegeben.

3.5 Genetische Verfahren

Die Grundidee der genetischen Verfahren ist der biologischen Evolution entlehnt. Zu einem Suchproblem wird eine Anfangspopulation von zufällig generierten Lösungen entworfen. Der Lösungsansatz jedes Individuums wird mittels einer Evaluierungsfunktion bewertet. Diejenigen Individuen, deren Lösungen vergleichsweise gut sind, vermehren sich. Bei der Vermehrung wird die Lösung der Eltern an die der Kinder in leicht mutierter Form weitergegeben. Die erfolgreichen Kandidaten reproduzieren

sich oft. Die Individuen, deren Bewertung durch die Evaluierungsfunktion schlecht ausfällt, sterben über einen gewissen Zeitraum hinweg aus. Durch die ständige zufällige Mutation und Selektion entstehen auf evolutionäre Weise nach mehreren Generationen Lösungen, die gut an ihre Evaluierungsfunktion angepaßt sind. Die so generierten Lösungen sind allerdings, je nach eingesetztem Zufallsgenerator, nicht zwingend reproduzierbar. Auch ist eine absolute Bewertung der Lösungen nicht möglich, da die optimale Lösung nach wie vor nicht bekannt ist. Bei diesem nicht erschöpfenden Verfahren läßt sich lediglich die relative Aussage treffen, daß eine Lösung besser ist als eine andere. Dies ist zugegebener Maßen nicht zufriedenstellend, aber die Methode ermöglicht das Finden suboptimaler Lösungen zu Problemen, die bisher als unlösbar galten. Folgendes Beispiel untermauert diese Behauptung.

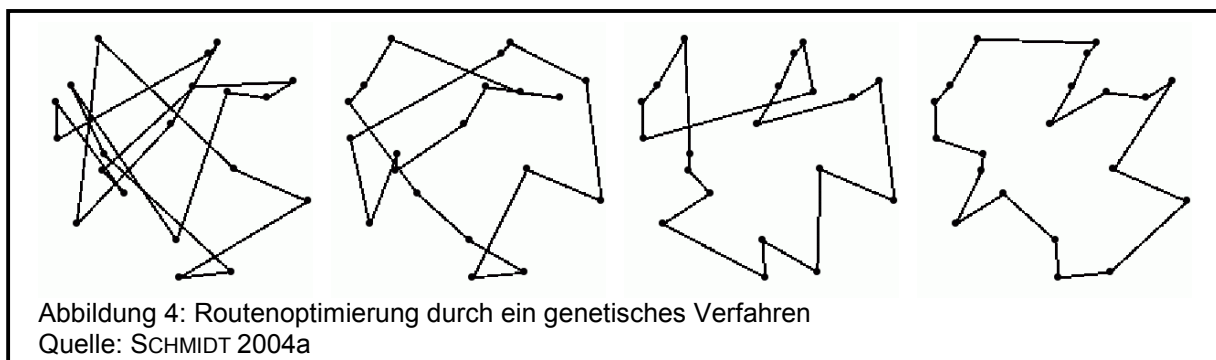
Anwendungsmöglichkeit

Sehr viele Probleme aus der Praxis lassen sich mit Hilfe evolutionärer Verfahren annäherungsweise lösen. Um die Leistungsfähigkeit des Verfahrens zu verdeutlichen, wird es an dieser Stelle auf das Problem des Handlungsreisenden angewandt. Ein Handlungsreisender muß Kunden in einer Reihe von Städten besuchen. Da er möglichst wenig unterwegs sein möchte, sucht er die kürzeste mögliche Route für seine Tour. Bei einer kleinen Anzahl von Städten N läßt sich das Problem noch mit Hilfe der blinden Suche lösen. Die Anzahl der möglichen Wege läßt sich nach der Formel

$$(N - 1)!$$

bestimmen. Für fünf Städte ergibt sich ein zu durchsuchender Suchraum von 24 Möglichkeiten, für 10 Städte bereits 362880 Möglichkeiten und für fünfzig Städte die unvorstellbare Anzahl von $6,08 \cdot 10^{62}$ Möglichkeiten. Bei der sehr optimistischen Annahme, daß pro Sekunde die Länge einer Milliarde Pfade berechnet werden könnte, würde die Suche nach dem kürzesten Pfad bei $N = 50$ immer noch $6,08 \cdot 10^{53}$ Jahre dauern. Diese Zahl entspricht in etwa knapp der Anzahl der Atome im Universum.

Zum Beginn der Suche wird eine Anfangspopulation von zufällig generierten Routen erzeugt (Abb. 4 links). Dazu wird die Reihenfolge, in der die Städte besucht werden, in einer Liste gespeichert. Für jedes Individuum wird die Länge seiner Route ermittelt. Diejenigen Kandidaten, deren Route relativ kurz ist, reproduzieren sich. Während des Reproduktionsvorgangs wird die Route zufällig leicht verändert, indem eine Stadt innerhalb der Liste zufällig verschoben wird. Diejenigen Individuen, die relativ schlechte, also lange Routen tragen, sterben aus. Nur die Erfolgreichen setzen sich durch. Nach dem Durchlaufen mehrerer Generationen wird eine Route gefunden, die sich gegenüber ihren Konkurrenten durchgesetzt hat und kürzer ist, als alle anderen bisher gefundenen Routen (Abb. 4, rechts). Ob diese Route optimal ist, läßt sich aufgrund der oben geschilderten Problematik nicht klären. Aber sie stellt eine gute, wenn vielleicht auch nur suboptimale Lösung des Problems dar.

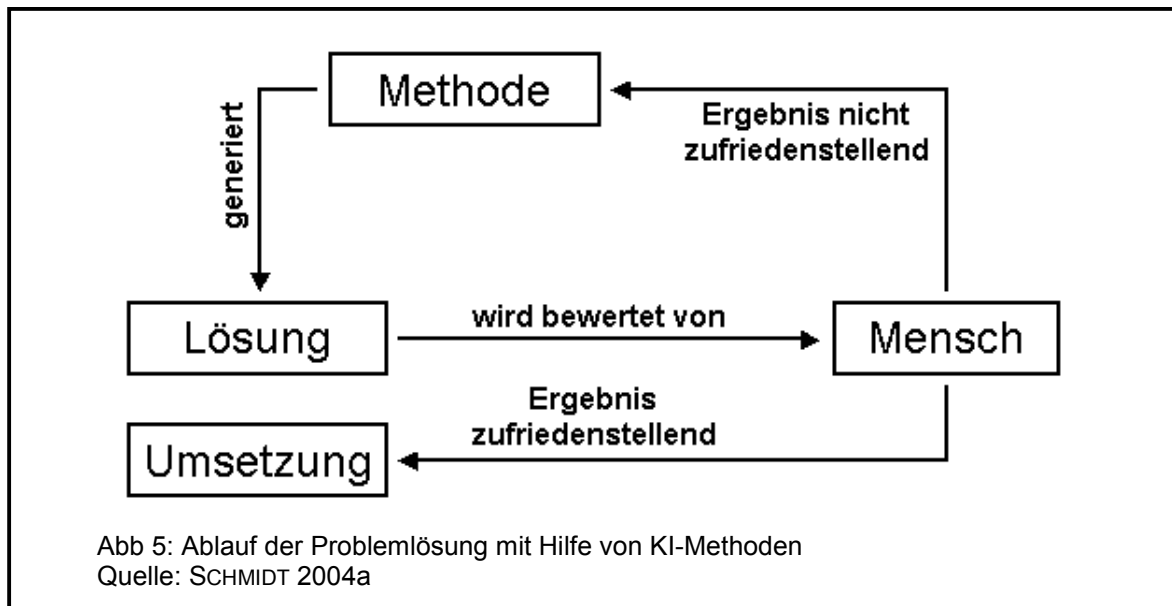


4 Fazit

Viele Ansätze der Geographie zur Lösung von Praxisproblemen leiden bisherige unter der Schwäche, nur bereits vorhandene Lösungsansätze bewerten zu können. Die gezeigten Beispiele haben, wenn auch auf abstrakter Ebene, veranschaulicht, daß Computer prinzipiell in der Lage sind, selbst Problemlösungen zu generieren. Methodisch bedingt sind die gefundenen Lösungen nicht immer absolut bewertbar, eventuell nicht einmal reproduzierbar. Diese Eigenschaft teilen sie mit Lösungsvorschlägen, die menschengemacht sind. Dies ist aus wissenschaftlicher Sicht unbefriedigend. Die Geographie steht jedoch in der Pflicht, Lösungen zu Problemen aus der Praxis anbieten zu können. Möglicherweise heiligt in diesem Fall der Zweck die Mittel. Die Erforschung der Anwendung von KI-Methoden in der Geographie steckt noch in den Kinderschuhen. Trotzdem wird an dieser Stelle die Ansicht vertreten, daß die klassischen Bewertungs- und Analysemethoden in Zukunft um die hier nur im Ansatz präsentierten Optimierungsverfahren erweitert werden können. Abb. 5 zeigt die Vorstellung

der Autoren, wie in nicht zu ferner Zukunft Praxisprobleme angegangen werden könnten. Die Aufgabe des Fachmanns besteht darin, sofern die Problemstellung dies ermöglicht, das Problem in eine computergerechte Form zu bringen und die Definition einer guten Lösung zu erstellen. Den Prozeß der Suche kann der Computer in Verbindung mit KI-Methoden vornehmen. Er tut dies unvoreingenommen und gemäß den Anweisungen, die er bekommt. Natürlich muß der Mensch immer die letzte Instanz bleiben, die eine Lösung auf ihre Plausibilität und Umsetzbarkeit prüft.

Schwierig ist derzeit noch die fehlende Anbindung von Simulations- oder Modellbildungssoftware an bestehende GI-Systeme. So lange die Softwarehersteller die erforderlichen Funktionalitäten nicht in ihre Programme einbinden, müssen die zu verarbeitenden Daten über spezielle Schnittstellen in eine externe, in der Regel selbst zu programmierende Software exportiert werden. Zur graphischen Darstellung der Ergebnisse ist ein Reimport der Daten in ein GI-System möglich.



Literatur

- OPENSHAW, S. & CH. OPENSHAW (1997): Artificial Intelligence in Geography. Chichester.
SCHMIDT, Th. (2004a): Methoden der Künstlichen Intelligenz in der Angewandten Geographie. Dipl.-Arb., Univ. Tübingen, Inst. f. Geogr. (unveröffentlicht)
WINSTON, P. H. (1992): Artificial Intelligence. – Boston.

Außerdem wurden verwendet:

- SCHMIDT, Th. (2004b): <http://www.geosimulation.de>
WILENSKY, U. (1999): NetLogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL, <http://ccl.northwestern.edu/netlogo>

Abbildungsunterschriften:

- Abb. 1: Ablauf einer Problemlösung mit bisherigen Verfahren
Abb. 2: Erläuterung siehe Text
Abb. 3: Erläuterung siehe Text
Abb. 4: Routenoptimierung durch ein genetisches Verfahren
Abb. 5: Ablauf der Problemlösung mit Hilfe von KI-Methoden

Dr. Hans-Joachim Rosner
Thorsten Schmidt
Universität Tübingen
Geographisches Institut
Rümelinstraße 19-21
72074 Tübingen
hans-joachim.rosner@web.de
Tho_Schmidt@web.de