

Verhaltensbasierte Ansätze in der KI —
Rodney A. Brooks

Thilo Weigel

24. Juni 1997

Seminar: Zielmanagement und Handlungskontrolle in autonomen Systemen
Abteilung für Grundlagen der Künstlichen Intelligenz
Institut für Informatik
Universität Freiburg
Sommersemester 1997

1 Einleitung

Diese Arbeit stellt die schriftliche Ausarbeitung eines Vortrages dar, der im Rahmen des Hauptseminars *Zielmanagement und Handlungskontrolle in autonomen Systemen* im Sommersemester '97 gehalten wurde. Sie basiert auf den Artikeln *Intelligence Without Reason* [1] und *A Robust Layered Control System for a Mobile Robot* [2] von Rodney A. Brooks.

Das zentrale Forschungsinteresse von Brooks gilt dem Gebiet der Robotik. In seiner Arbeit mit autonomen, mobilen Robotern beschäftigt er sich vor allem mit Strategien, wie für Roboter auf niedriger Ebene ein robustes Verhalten implementiert werden kann. Brooks favorisiert hierfür einen verhaltensbasierten Ansatz, bei dem komplexe Funktionen des Roboters nicht explizit programmiert werden, sondern sich aus dem Zusammen- und Gegen-spiel einzelner Module ergeben sollen.

Brooks kritisiert die Ansätze der traditionellen KI in vielen Punkten und macht deutlich, warum er seinen Ansatz als den geeigneteren ansieht, künstliche intelligente Systeme zu entwickeln, die unter Bedingungen der „realen Welt“ mit robustem „Verhalten“ ihre Aufgaben erledigen.

Auf einige Kritikpunkte, zu denen der Brooks'sche Ansatz besonders in Kontrast steht, soll im folgenden Abschnitt 2 zusammenfassend eingegangen werden. Abschnitt 3 soll dann eine Übersicht über die theoretischen Überlegungen dieses Ansatzes geben und Abschnitt 4 die konkrete Architektur vorstellen.

2 Traditionelle KI

In dem Artikel *Intelligence Without Reason* [1] unterzieht Brooks die gesamte KI-Forschung seit der Dartmouth-Konferenz im Jahre 1956 einer kritischen Betrachtung. Er stellt die These auf, daß mit ihr eine Festlegung auf Prinzipien stattfand, die die Forschung bis heute stark beeinflussen. Brooks sieht insbesondere eine Festlegung auf die Vorstellung eines rationalen Agenten, der anhand einer internen Repräsentation durch Suche einen Plan für sein späteres Handeln erstellt.

Die Annahme auf dem Gebiet der Robotik, die Sensortechnologie würde eines Tages so weit ausgereift sein, daß „ideale“ Sensordaten zur Verfügung stehen werden, stellt für Brooks nur ein Beispiel dar, wie schon früh, unzulässigerweise, vereinfachende Grundannahmen gemacht und einschränkende Arbeitshypothesen aufgestellt wurden. Diese seien zudem mit der Zeit in Vergessenheit geraten, so daß er vielen Forschungsarbeiten, die zu einem gewissen Grad auf solchen Annahmen gründen, fehlenden Bezug zur Realität

attestiert.

„The founders of the bodies of work are quite aware of these assumptions, but over time as new people come into the fields, these assumptions get lost, forgotten, or buried, and the work takes on a life for its own sake.“ [1, S. 3]

In den folgenden Abschnitten soll nun detaillierter auf einige Kritikpunkte eingegangen werden.

2.1 Computer

Eine wichtige These Brooks'scher Kritik bezieht sich auf den Zusammenhang, den er zwischen zur Verfügung stehender Computerarchitektur und den Leitideen der KI-Forschung sieht. Die in den letzten Jahrzehnten dominierende von-Neumann-Architektur hat laut Brooks nicht nur die Wahl der zu betrachtenden Probleme beeinflusst, sondern auch die zu ihrer Lösung entwickelten Algorithmen:

„... the architecture of our computers influences our choice of problems on which to work, our models of thought and our algorithms ...“ [1, S. 3]

Brooks stellt fest, daß selbst das Bild, das sich Forscher von der Funktionsweise des menschlichen Verstandes machten, von der Computerarchitektur beeinflusst wurde. Dies führte seiner Meinung nach zu Ansätzen, die auf keinen Fall tauglich sind, menschliche Intelligenz zu erklären oder nachzubilden. Als Beispiel führt Brooks das Schachspiel an, von dem Turing noch glaubte, daß es der Erforschung des menschlichen Geistes dienlich sein würde. Deep Blue lehrt uns heutzutage, daß das Resultat aufwendiger Berechnungen zwar den Resultaten menschlicher Entscheidungsfindung überlegen sein kann, damit aber nichts zur Erforschung des menschlichen Geistes beigetragen worden ist.

2.2 Roboter

Brooks glaubt nicht, daß zukünftige Sensoren einem Roboter einmal ein komplettes Bild der Welt vermitteln werden können. Er hält daher Ansätze, die allein auf der Basis eines internen Weltmodells Handlungen planen, für schlichtweg unrealistisch.

„The key problem that I see with all this work is that it relied on the assumption that a complete world model could be built internally and then manipulated. . . eventually computer vision will catch up and provide such world models — I don't believe this based on the biological evidence. . . or complete objective models of reality are unrealistic and hence the methods of Artificial Intelligence that rely on such models are unrealistic.“ [1, S.9]

Brooks hebt hervor, daß selbst der prominente Roboter *Shakey*, der mit einer internen Repräsentation arbeitete und durchaus komplexe Aufgaben bewältigen konnte in der „realen Welt“ keinerlei Erfolg gehabt hätte. Er operierte lediglich in einer Miniwelt mit Objekten, deren Form und Farbe speziell auf die Kameras und Erkennungsalgorithmen abgestimmt waren.

2.3 Biologie

Brooks weist darauf hin, daß sich die Organisationsstrukturen des menschlichen Gehirns von denen der traditionellen KI-Ansätze erheblich unterscheiden. Im Gegensatz zum Modell eines rationalen Agenten, der auf einer Eingabe basierend eine Berechnung durchführt und diese dann ausgibt, charakterisiert er die Funktionsweise des Gehirns als „a mess of many mechanisms working in various ways“. Brooks erwartet nicht, daß all diese Mechanismen eines Tages wirklich verstanden werden können, vermißt aber dennoch, daß nicht einzelne Hinweise, wie zum Beispiel das Prinzip der Hormone, in die Forschung integriert werden.

Am Beispiel optischer Täuschungen macht Brooks deutlich, daß der sensorische Apparat des Menschen bei weitem nicht so perfekt ist, wie es von den Sensoren zukünftiger Roboter erwartet wird. Die Natur stellt aber offensichtlich einen Mechanismus bereit, der die Schwächen des sensorischen Apparates kompensiert.

Brooks sieht die den bisherigen Ansätzen zugrundeliegenden Prinzipien nicht unbedingt als geeignet an, künstliche intelligente Systeme zu entwickeln. Humanbiologische Betrachtungen machen für ihn zudem deutlich, daß diese Prinzipien keineswegs notwendige Voraussetzung für Intelligenz sind:

„What is the point of all this? The traditional Artificial Intelligence model of representation and organization along centralized lines is not how people are built. Traditional Artificial Intelligence methods are certainly not necessary for intelligence then, and so far they have not really been demonstrated to be sufficient in situated, embodied systems.“ [1, S. 13]

3 Der Ansatz von Brooks

Der Brooks'sche Ansatz basiert auf autonomen, mobilen Robotern. Diese sollen in beliebigen Kontexten der „realen Welt“ funktionstüchtig sein und dabei trotz unvollständiger Sensordaten und sich ändernder Umweltbedingungen ein robustes „Verhalten“ zeigen. Brooks gibt sich bezüglich seines Ansatzes betont optimistisch:

„I think that the new approach can be extended to cover the whole story, both with regards to building intelligent systems and to understanding human intelligence — the two principal goals identified for Artificial Intelligence at the beginning of the paper.“
[1, S. 16]

Zunächst sollen für den Ansatz zentrale Begriffe geklärt werden.

3.1 Schlüsselbegriffe

3.1.1 Intelligenz

Brooks begreift *Intelligenz* als ein fast allem menschlichen Handeln und Denken innewohnendes Merkmal. Informal beschreibt er Intelligenz daher als „the sort of stuff that humans do, pretty much all the time“. Er weist darauf hin, daß keine klare Trennlinie zwischen intelligentem und reaktivem Verhalten existiert. Die Struktur der Intelligenz sei immer auch durch die Umwelt bestimmt. In der physikalischen Verbindung zur Welt, in der Organisation der „Sensoren“ oder der Struktur elementarer Prozesse sei damit schon die Basis oder vielmehr die notwendige Voraussetzung für Intelligenz zu sehen.

„... but the source of intelligence is not limited to just the computational engine. It also comes from the situation in the world, the signal transformations within the sensors, and the physical coupling of the robot with the world.“ [1, S. 3]

Brooks sieht es als unerlässlich an, motorische Fähigkeiten und grundlegende Verhaltensmuster mehr in den Mittelpunkt des Forschungsinteresses zu rücken. Diese seien in längerem Evolutionszeitraum erworben worden als der „High-Level-Intellekt“ und verdienten daher auch mehr Beachtung bei der Grundlagenforschung.

3.1.2 Emergenz

Traditionelle Ansätze der KI versuchten bisher stets, ein System mit „Intelligenz“ auszustatten, indem sie vorgaben, wie in welcher Situation zu reagieren sei. Brooks hält dem entgegen, daß wirkliche Intelligenz einem System nicht a priori gegeben sein könne, sondern vielmehr aus der Interaktion mit der Welt und der Interaktion der Komponenten des Systems untereinander entstehe.

Konsequenterweise verfolgt Brooks in seinem Ansatz die Strategie, einzelne Module zu implementieren, von denen jedes für sich in gewissem Sinne verhaltenproduzierend ist. Ohne vorher explizit vorgegeben worden zu sein, soll dann, lediglich durch das Zusammenspiel der Module, ein Gesamtverhalten entstehen, das mehr als nur die Summe der Einzelverhalten darstellt.

„The intelligence of the system emerges from the system’s interaction with the world and from sometimes indirect interactions between its components — it is sometimes hard to point to one event or place within the system and say that is why some external action was manifested.“ [1, S. 3]

„Thought and consciousness will not need to be programmed in. They will emerge.“ [1, S. 22]

3.1.3 Situationsbezogenheit (Situativeness)

Der kontinuierliche Bezug zu den aktuellen Umweltgegebenheiten ist laut Brooks für einen autonomen, mobilen Roboter unabdingbar. Aktionen sollten weniger auf Basis von internen, abstrakten Repräsentationen und Plänen geschehen, sondern vielmehr dauernd von den aktuellen Sensordaten abhängig gemacht werden. Nur ein hochgradig reaktives System kann laut Brooks ein robustes Verhalten in komplexen, dynamischen Umgebungen gewährleisten.

„The robots are situated in the world — they do not deal with abstract descriptions but with the here and now of the world directly influencing the behavior of the system.“ [1, S. 3]

Als zu realitätsfern beschreibt Brooks daher Ansätze, die von einer idealen, statischen Umgebung ausgehen, Brooks zieht es vor, auf kein internes Weltmodell zurückzugreifen, denn „the world is our best model“ [1, S. 15].

3.1.4 Verkörperung (Embodiment)

Laut Brooks bekommt ein künstliches System erst mit einem „Körper“, durch den es mit der Welt interagiert, einen Sinn. Erst ein konkret vorliegender

Apparat stellt einen Bezug zur Umwelt her und gibt Informationen über die Welt und Aktionen eine Bedeutung. Für Brooks existieren sogar viele Anzeichen dafür, daß ein Gehirn ohne einen Körper überhaupt nicht denkbar ist. Er folgert, daß für KI-Forschung, die sich an ihrem tatsächlichen Erfolg in der „realen Welt“ messen lassen will, Roboter das ideale Medium darstellen.

„The robots have bodies and experience the world directly — their actions are part of a dynamic with the world and have immediate feedback on their own sensations.“

3.2 Designprinzipien

Erklärtes Ziel des Ansatzes von Brooks ist die Entwicklung eines autonomen, intelligenten Agenten, der ohne Einschränkungen in natürlichen, dynamischen Umgebungen operieren kann. Der Agent soll als Roboter „verkörpert“ stets „situationsbezogen“ operieren. Das Verhalten des Roboters soll robust sein, d.h. Fehler oder Ausfall der Sensoren oder Änderungen der Umwelt sollen zumindest ein der Situation angemessenes Verhalten hervorrufen.

Der Ansatz geht davon aus, daß komplexes Verhalten nicht unbedingt ein Produkt eines komplexen Kontrollsystems sein muß. Er baut auf ein dezentrales Netzwerk verschiedener Module, die für sich allein aktive Recheneinheiten darstellen. Die Zustände eines Moduls und deren Variablen stellen die einzige Form eines Speichers für das System dar. Es sind weder ein zentrales Modell der Welt, noch explizite Repräsentationsmechanismen, Problemlösen oder eine zentrale Kontrolleinheit explizit vorgesehen.

3.3 Subsumptionsarchitektur

Herkömmliche Kontrollsysteme wurden meist nach dem in Abbildung 1 dargestellten Prinzip organisiert.

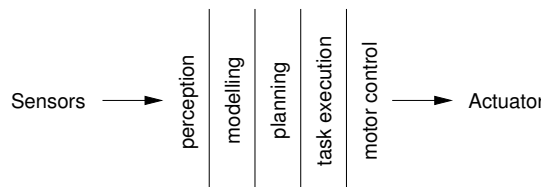


Abbildung 1: Traditionelle Dekomposition eines Kontrollsystems in funktionale Module

Informationen durchliefen einen Zyklus, der stets vollständig durchlaufen werden mußte, um aufgrund einer neuen Sensorinformation eine Aktion auszuführen. Eine horizontale Aufteilung in Teilprobleme hat zur Konsequenz, daß das System nur dann funktionsfähig ist, wenn alle Teilprobleme des Zyklus implementiert sind und fehlerfrei arbeiten. Nachteile dieses Ansatzes sind schlechte Erweiterbarkeit und hohe Fehleranfälligkeit, da bei Ausfall eines Moduls der gesamte Zyklus unterbrochen ist. In dem von Brooks vorgeschlagenen Kontrollsystem soll daher das Teilproblem wie in Abbildung 2 vertikal aufgeteilt werden.

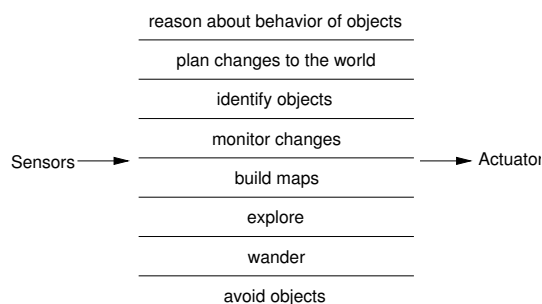


Abbildung 2: Dekomposition eines Kontrollsystem auf Grundlage aufgabenerfüllenden Verhaltens

Das System ist aus verschiedenen Schichten aufgebaut, wobei höhere Schichten zwar auf niedrigeren aufbauen, das System aber auch schon mit nur einer Schicht funktionsfähig ist. Mit jeder Schicht gewinnt das System an Komplexität. Während die unterste Schicht lediglich gewährleisten soll, daß Hindernissen ausgewichen wird, soll die höchste Schicht komplexe Schlüsse über die Umwelt ziehen und Pläne modifizieren können. Brooks unterscheidet insgesamt acht Schichten unterschiedlicher „Kompetenz“.

1. Weiche Hindernissen aus.
2. Fahre ziellos umher.
3. Erforsche die Umgebung.
4. Erstelle eine Karte und plane Wege.
5. Stelle Änderungen in der „statischen“ Umgebung fest.
6. Schlußfolgere über Objekte und führe mit ihnen Aktionen aus.

7. Formuliere und führe Pläne aus, die auch den Zustand der Welt verändern können.
8. Schlußfolgere über die Welt und modifiziere die Pläne entsprechend.

Jede der vorher skizzierten Kompetenzschichten entspricht einer Schicht des Kontrollsystems. Diese Schichten sind unabhängig voneinander gebaut, höhere Schichten können aber Daten von niedrigeren Schichten untersuchen, oder niedrigeren Schichten Daten zuführen und damit deren eigentliche Eingabe unterdrücken. Eine Schicht subsumiert damit eine niedrigere.

3.4 Endliche Automaten

Eine Kontrollschicht setzt sich aus einzelnen wie in Abbildung 3 gezeigten Modulen zusammen.

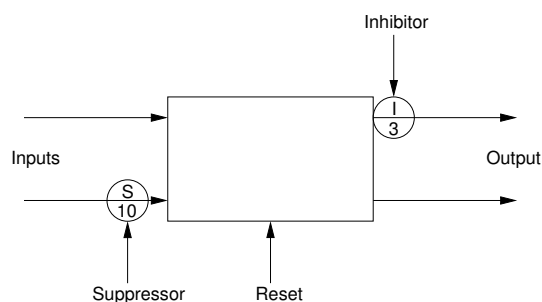


Abbildung 3: Ein Modul mit Ein- und Ausgabeleitungen. Eingangssignale können unterdrückt werden und durch das unterdrückende Signal ersetzt werden. Die Ausgabe eines Moduls kann verhindert werden

Prinzipiell kann ein Modul als ein endlicher Automat angesehen werden, der folgende Zustände annehmen kann:

Ausgabe Eine Ausgabenachricht, berechnet als eine Funktion der Eingabepuffer und der Modulvariablen, wird zur Ausgabelitung gesendet. Das Modul nimmt einen neuen Zustand an.

Seiteneffekt Eine der Variablen des Moduls wird neu gesetzt, das Modul nimmt einen neuen Zustand an.

Bedingte Verzweigung Ein Prädikat, abhängig von den Variablen des Moduls und des Eingabepuffers, wird berechnet. Abhängig vom Resultat nimmt das Modul einen von zwei Zuständen an.

Ereignisverzweigung Eine Abfolge von Zuständen und Bedingungen werden überwacht, bis eines der Ereignisse wahr wird.

NIL Bezeichnet den Startzustand, der jederzeit durch ein „Reset“ wieder hergestellt werden kann.

Über Ein- und Ausgabeleitungen kommuniziert ein Modul mit anderen, wobei ein einelementiger Puffer einer Eingabeleitung dem Modul jederzeit die letzte Eingabe bereitstellt. Nachrichten können verlorengehen, wenn eine neuankommende eine alte Nachricht überschreibt. Jeweils während eines gewissen Zeitraumes kann die Ein- oder Ausgabe eines Moduls unterdrückt werden. Die Ausgabe eines Moduls kann jeweils die Eingabe für ein anderes Modul liefern, die Ausgabe eines anderen Moduls unterdrücken oder sogar die Eingabe eines anderen Moduls durch die eigene Ausgabe ersetzen.

4 Ein Kontrollsystem für einen mobilen Roboter

Hier soll die konkrete Implementierung vorgestellt werden, für die in den letzten Abschnitten die theoretischen und „philosophischen“ Grundlagen dargestellt wurde. Die Implementation ist in dem Artikel *A Robust Layered Control System for a Mobile Robot* [2] beschrieben.

4.1 Die erste Kontrollschicht

Die Kontrollebene der untersten Stufe soll lediglich gewährleisten, daß der Roboter Hindernissen ausweicht oder sie meidet. Der Aufbau ist in Abbildung 4 dargestellt.

Ausgehend von den durch das *sonar*-Modul bereitgestellten Sonardaten überprüft das *collide*-Modul, ob sich direkt vor dem Roboter ein Objekt befindet. Ist dies der Fall, so sendet es einen *Halt*-Befehl an das *forward*-Modul, das dies in Form von Motorsteuerungssignalen dem Roboter übermittelt. Parallel dazu errechnet das *feelforce*-Modul aus den Positionen aller wahrgenommenen Objekte eine abstoßende Kraft, weg von allen Objekten. Das *runaway*-Modul überwacht den Wert dieser Kraft. Überschreitet sie einen bestimmten Schwellenwert, so wird die Information an die Module *turn* und *forward* weitergeleitet, wo die Aktionen in konkrete Motorbefehle umgesetzt werden. Allein die Taktik, vor Hindernissen zu stoppen und vor sich nähernden Objekten „davonzulaufen“ soll somit gewährleisten, nie mit einem Objekt zusammenzustoßen.

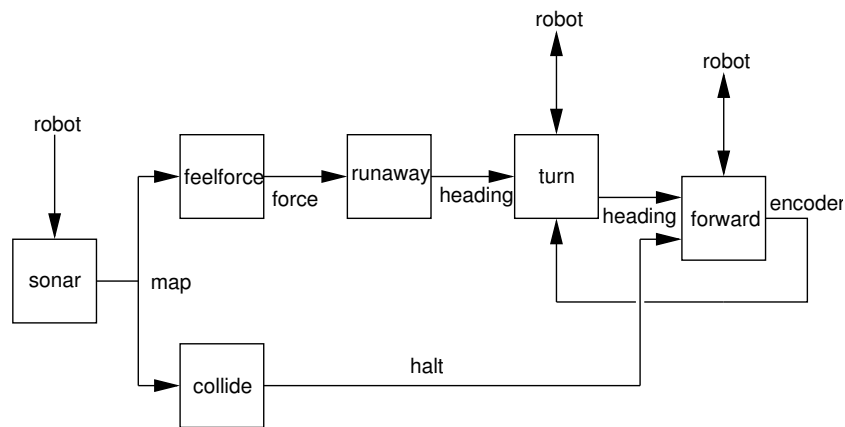


Abbildung 4: Kontrollschicht 1

4.2 Die zweite Kontrollschicht

Die zweite Kontrollschicht soll den Roboter ziellos umherwandern lassen, wobei die erste Schicht nach wie vor verhindern soll, daß der Roboter mit Hindernissen zusammenstößt. Abbildung 5 zeigt die nun um Schicht 2 erweiterte Kontrollschicht.

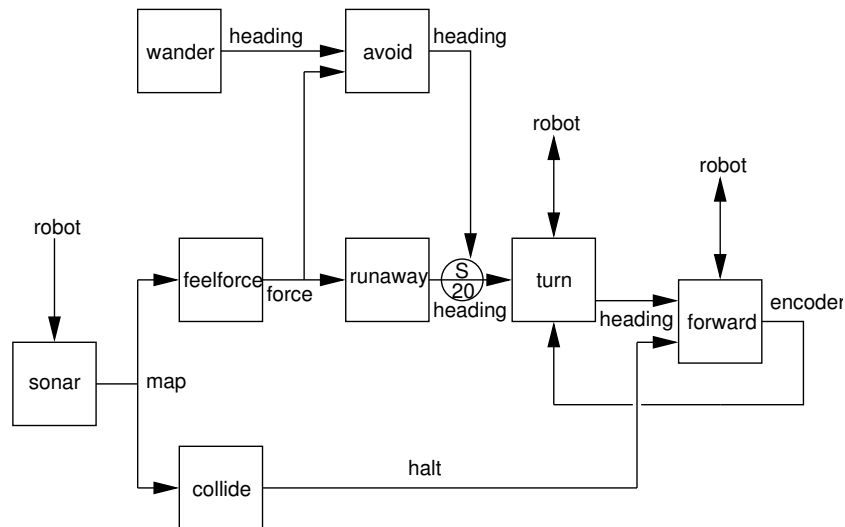


Abbildung 5: Die um die zweite Schicht erweiterte Kontrollstruktur

Das neu hinzugekommene *wander*-Modul erzeugt nun in gewissem Sekundentakt eine neue Richtung für den Roboter. Das *avoid*-Modul kombiniert

diese Richtungsangabe mit der des *feelforce*-Moduls und teilt die neue Richtungsinformation dem *turn*-Modul mit. Die Ausgabe des *runaway*-Moduls der untersten Stufe wird damit von dem *avoid*-Modul unterdrückt, die Berechnung des *runaway*-Moduls implizit von der des *avoid*-Moduls subsumiert.

4.3 Die dritte Kontrollschicht

Die dritte Kontrollschicht soll dem System ein exploratives Verhalten hinzufügen. Die Verbindung zu den ersten beiden Schichten wird in Abbildung 6 dargestellt.

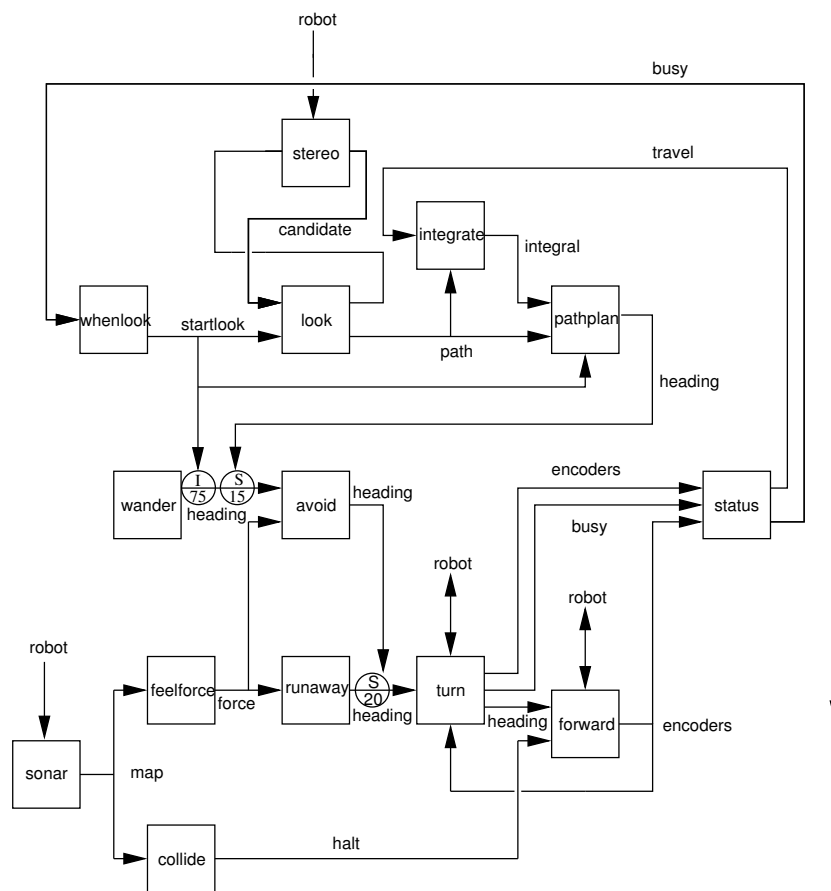


Abbildung 6: Die um die dritte Schicht erweiterte Kontrollstruktur

Ein *status*-Modul überwacht nun, ob der Roboter sich in Bewegung befindet. Ist dies nicht der Fall, so verhindert das *whenlook*-Modul die Ausgabe des *wander*-Moduls, um genügend Zeit zu haben, mit einer Kamera Bilder aufzunehmen und diese durch das *look*-Modul analysieren zu lassen. Der Prozeß der

Bildauswertung soll in zukünftigen Versionen noch durch Stereokamerabilder des *stereo*-Moduls erweitert werden. Das Ergebnis der Bildauswertung wird nun dem *pathplan*-Modul zugeführt, das das Ergebnis seiner Wegplanung dem *avoid*-Modul zuführt und dabei zielloses Umherwandern unterdrückt. Das *integrate*-Modul überwacht schließlich anhand der ihm zugeführten Informationen bezüglich der Bewegungen des Roboters den Fortschritt bei der Bewegung zu einem Ziel hin und kann die weitere Wegplanung beeinflussen.

5 Diskussion

Einleitend stellte Brooks als die beiden wesentlichen Ziele der KI die Erforschung menschlicher Intelligenz und die Entwicklung künstlicher intelligenter Systeme dar. Verhaltensbasierte, „verkörperte“ Roboter hält er für am weitaus besten geeignet, zu beiden Zielrichtungen einen Beitrag zu leisten.

Der Ansatz von Brooks verfügt über einige interessante Aspekte. Eine engere Anlehnung an das Vorbild der Biologie etwa oder die Idee der Emergenz von Intelligenz erscheinen vielversprechende Konzepte zu sein. Dennoch wirft der Ansatz auch viele kritische Fragen auf, die an der beliebigen Erweiterbarkeit des Systems zweifeln lassen.

Die Arbeit mit dem verhaltensbasierten Saphira-Kontrollsystem für die Roboter der Abteilung Künstliche Intelligenz hat bisher gezeigt, daß die Koordination von einzelnen Verhaltensbausteinen ein nicht triviales Unterfangen ist, da konkurrierende Verhaltensweisen sich nur allzuoft zum Stillstand des Roboters summieren.

Das Erscheinungsdatum der Artikel liegt schon einige Jahre zurück. Die Tatsache, daß seither nichts von der Vollendung weiterer „Kompetenzschichten“ bekannt geworden ist, verleitet zu der Vermutung, daß sich der Komplexitätsgrad der Architektur nicht ohne weiteres steigern ließ. Es ist schwer vorstellbar, daß ein künstliches System ganz ohne zentralen Speicher, symbolische Repräsentation und Planung komplexe künstlich intelligente Leistungen vollbringen kann.

Der Ansatz von Brooks, einhergehend mit seiner provokanten Kritik an der traditionellen KI stellt sicherlich einen interessanten Diskussionsgegenstand dar. Bisher scheint er die KI-Forschung dennoch noch nicht revolutioniert zu haben.

Literatur

- [1] Brooks, R. A. (1986). Intelligence Without Reason. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)*, Sydney.
- [2] Brooks, R. A. (1986). A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation* RA-2(1):14-23.