

Cognitive Humanoid Robotics: Steps toward Learning and Behavior Organization

Christian Goerick, Edgar Körner

2010

Preprint:

This is an accepted article published in Automatisierungstechnik. The final authenticated version is available online at: [https://doi.org/\[DOI not available\]](https://doi.org/[DOI not available])

Kognitive Humanoide Robotik: Entwicklungsschritte in Lernen und Verhaltensorganisation

Cognitive Humanoid Robotics: Steps Toward Learning and Behavior Organization

Christian Goerick, Edgar Körner, Honda Research Institute Europe

Zusammenfassung In diesem Beitrag geben wir einen Überblick über unsere Forschungsarbeiten auf dem Gebiet der humanoiden Robotik mit dem Schwerpunkt auf kognitiven Systemen. Insbesondere betrachten wir eine Serie von Systemen, die wir gemäß unseres hierarchischen inkrementellen System-

ansatzes erforscht, realisiert und analysiert haben. ▶▶▶

Summary In this paper we review our work concerning cognitive robotics. We present a series of systems we have been researching, implementing and analyzing according to our incremental hierarchical system approach.

Schlagwörter Kognitive Robotik, hierarchische Verhaltenskontrollsysteme, Semantik lernen, Mensch-Roboter-Interaktion ▶▶▶

Keywords Cognitive robotics, hierarchical behavior control systems, semantics learning, human robot interaction

1 Einleitung

Die humanoide Robotik ist durch zwei gekoppelte Ansätze charakterisiert: Der erste als konstruktiver, mit dem Ziel ein Artefakt zu kreieren, das menschenähnliche Fähigkeiten zeigt, mit einer Bewertung auf der extern beobachtbaren Ebene. Der zweite als erkenntnistheoretischer, mit dem Ziel zu verstehen, wie die internen Mechanismen organisiert werden können, um solcherlei Leistungen zu erzeugen, mit einer Bewertung auf der internen, funktionalen Ebene. Gekoppelt werden sie durch die Notwendigkeit des Verständnisses der internen Mechanismen beim ersten und der Überprüfung der gewonnen Erkenntnisse mit Hilfe von implementierten Systemen beim zweiten Ansatz. Bei der Wahl der internen Mechanismen liegt die Anlehnung an die Biologie und die Evolution auf der Hand, da sie das Vorbild dieser Forschungsrichtung hervorgebracht haben. Besondere Rollen kommen hier dem humanoiden Körper und dem menschlichen Gehirn zu. Ein humanoider Körper erlaubt durch die Ähnlichkeit zum menschlichen Körper prinzipiell eine Ähnlichkeit bei der Konzeptbildung. Ausserdem

ist er besonders gut geeignet in antropozentrischen Umgebungen den Menschen zu unterstützen. Das Gehirn ist das einzige existierende kognitive System, von dessen Verständnis wir uns große Fortschritte für die praktische Intelligenz von Artefakten erwarten.

In diesem Artikel werden wir unsere aktuellen Arbeiten zu humanoiden Robotern darstellen, wobei unser Schwerpunkt auf der schrittweisen Entwicklung der mentalen Fähigkeiten liegen wird. Allgemein verwandte Arbeiten sind [1–3] und [4], detailliertere werden an den entsprechenden Stellen direkt diskutiert. Einen Überblick zu kognitiven Architekturen findet sich in [5].

Wir verfolgen einen systemorientierten hierarchischen Ansatz, mit dem wir uns schrittweise höheren kognitiven Leistungen nähern. Systemorientiert bezieht sich auf eine entsprechende Konzeption des Systems als Ganzes, da wir der Meinung sind, dass die angestrebten Fähigkeiten nur aus einem wohl verstandenen Zusammenspiel der einzelnen Teile entstehen können.

Zur Beschreibung unserer Systeme haben wir SYSTEMATICA entwickelt, ein Modellierung zur Beschrei-

bung hierarchischer verhaltensgenerierender Systeme [6]. Damit können sowohl einzelne Systeme als auch unterschiedliche Systemklassen einheitlich beschrieben und miteinander verglichen werden.

2 Reaktive Kontrolle und visuelles Lernen: ALIS 1

Das erste erforschte und implementierte System trägt den Namen ALIS 1, eine Abkürzung für *Autonomous Learning and Interacting System*. Es realisiert eine reaktive autonome Verhaltenskontrolle auf der Plattform des humanoiden Roboters von Honda. Die Architektur des Systems ist in Bild 1 dargestellt. Das System ist in zwei Teile aufgeteilt, die jeweils auf dem Roboter oder auf stationären Computern ablaufen. Auf dem Roboter (Schicht 1) sind alle für die direkte Sicherheit und Stabilität des Roboters verantwortlichen Komponenten verortet. Dies sind die Servokontrolleinheiten, die Selbstkollisionserkennung und -vermeidung, die Ganzkörperkontrolle und die Trajektoriengenerierung. Weiterhin werden an Bord des Roboters die visuellen und auditorischen Signale gemessen. Die Ansteuerung des Roboters erfolgt mit Ausnahme der untersten Regelung nicht mit Hilfe von Gelenkwinkeln sondern im Aufgabenraum. Dies erlaubt eine effiziente und abstrakte Kontrolle für darauf aufbauende Schichten. Die darüber liegende Schicht 2 realisiert eine auf visueller Salienz basierende Blickrichtungsteuerung, Schicht 3 berechnet Blickrichtungen auf Grund von auditorischer Salienz, die höher priorisiert sind als die visuellen, aber mit diesen in der Schicht 2 fusioniert werden können. Damit ist es

für einen menschlichen Interaktor möglich durch allgemeine visuelle oder auditorische Reize die Blickrichtung des Roboters in eine bestimmte Richtung zu lenken und damit in Interaktion zu treten. Die Schicht 4 realisiert Extraktion und Fixation so genannter Proto-Objekte. Dabei handelt es sich um grundlegende allgemeine Objekte im Raum mit unbekannter Identität, deren Ort und Lage aber verfolgt werden kann. In der vorliegenden Implementierung sind dies alle Objekte innerhalb des so genannten Peripersonal-Raumes um den Roboter herum. Der Peripersonal-Raum entspricht dem ohne Fortbewegung durch den Roboter zu manipulierenden Raum. Bei Studien mit kleinen Kindern hat sich gezeigt, dass dies der Raum ist, in dem primär Informationen aufgenommen und gelernt werden. Von den Objekten innerhalb des Peripersonal-Raumes wird eines ausgewählt und fixiert. Die visuelle Erscheinung eines solchermassen fixierten Proto-Objektes kann in Schicht 5 gelernt und mit einem Namen aus einem vordefinierten Reservoir belegt werden. Dieses Lernen erfolgt während des laufenden Betriebs und kann direkt zur Erkennung des gerade fixierten Proto-Objektes genutzt werden. Die Assoziation zwischen dem fixierten Proto-Objekt und seiner visuellen Identität für die momentan vom Roboter beobachtete Szene erfolgt in Schicht 6. Schicht 7 beinhaltet die Evaluation der Szene und die Verhaltensgenerierung des Körpers. Hier wird determiniert dass der Roboter auf ein Objekt zeigt oder dass mit Hilfe von Laufbewegungen der Abstand zum fixierten Proto-Objekt innerhalb eines bestimmten Interaktionsintervalls gehalten wird. Schicht 8 berechnet aus Sprachsignalen des Interaktors entsprechende interne

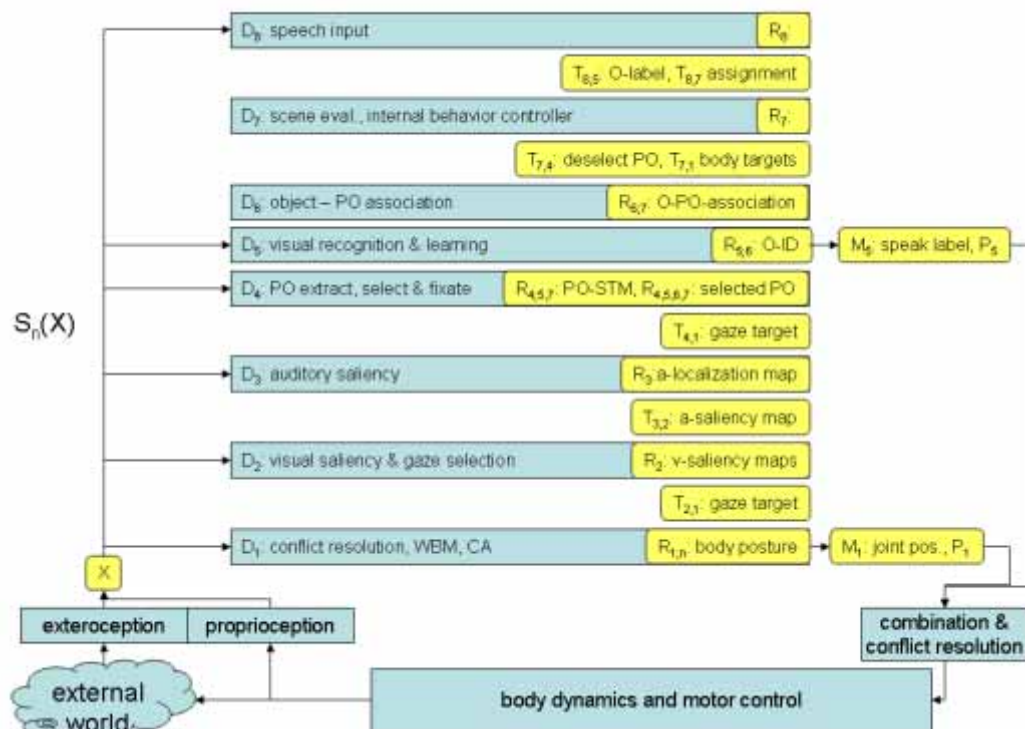


Bild 1 Schematische Darstellung von ALIS 1.

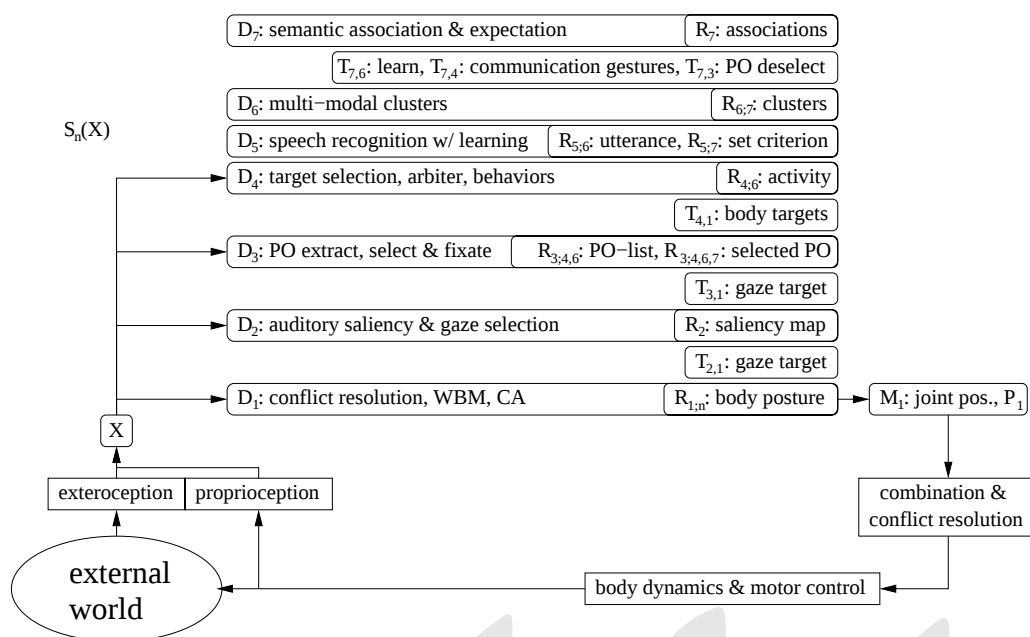


Bild 2 Schematische Darstellung von ALIS 2.

Signale bezüglich der Lernkontrolle und des auszuwählenden Namens. Die Schichten sind jeweils autonom und bauen nur von unten nach oben aufeinander auf. Untere Schichten realisieren autonome Teilverhalten und stellen berechnete Repräsentation für darüber liegende Schichten zur Verfügung, in Bild 1 mit R bezeichnet. Höhere Schichten können mit Hilfe von Top-Down-Signalen auf darunter liegende einwirken, in Bild 1 mit T bezeichnet. Eine detaillierte Beschreibung des Systems findet sich in [6].

Eine Interaktion mit ALIS 1 ist in Bild 3 illustriert und stellt sich wie folgt dar. Ohne einen Interaktionspartner fixiert der Roboter getrieben durch Saliens verschiedene Punkte in seiner Umgebung. Nähert sich ein Interaktionspartner und macht sich durch auditorische Signale, wie zum Beispiel Rufen, bemerkbar, so wird der Roboter versuchen die auditorische Quelle zu fixieren, was den Interaktionspartner in das visuelle Feld des Roboters bringt. Wird dann dem Roboter in den Peripersonal-Raum hinein ein Gegenstand entgegengehalten, so wird dieser fixiert und der Abstand wird mittels Laufbewegungen möglichst konstant gehalten. Der fixierte Gegenstand wird auf Grund seiner visuellen Erscheinung klassifiziert oder bei entsprechenden Äusserungen des Interaktionspartners erlernt. Der Lerndialog umfasst die Benennung des Gegenstandes gemäß einer definierten Liste und die Korrektur bei falscher Klassifikation. Nach erfolgreicher Klassifikation wird, falls möglich, ein anderer, noch nicht betrachteter Gegenstand fixiert und klassifiziert. Alle während der Fixation gespeicherten Ansichten können für den Lernvorgang genutzt werden, auch wenn noch kein Lernkommando gegeben wurde. Damit wird die klassische Trennung zwischen Lernen und Erkennen aufgehoben. Das Lernen erfolgt für den Interaktionspartner ohne nen-

nenswerte Verzögerung, wobei das System robust genug ist um bei vollen Roboterbewegungen gut zu funktionieren. Das Erlernen von visuellen Ansichten auf einem humanoiden Roboter ist am direktesten mit den Arbeiten von Ude zu vergleichen, siehe dazu auch [7].

Zusammenfassend haben wir ein hierarchisch aufgebautes reaktives System, das sich automatisch in Interaktion begibt und darin verbleibt, und das innerhalb solcher Interaktionen visuell Objekte erlernen und erkennen kann. Dabei werden alle Bewegungen des Roboters in der Interaktion während der Laufzeit berechnet. Diese Freiheit bei der Bewegungsgenerierung wird durch die Selbstkollisionsvermeidung ermöglicht, die ein Zusammentreffen von Teilen des Roboters verhindert. Die Ganzkörperkontrolle ermöglicht harmonische



Bild 3 Interaktion des Roboters mit einem menschlichen Partner.

Bewegungen mit einer automatisierten Verteilung der Bewegungen auf die verschiedenen Freiheitsgrade. Die einzelnen Schichten innerhalb der Hierarchie arbeiten über eine lose Kopplung zusammen und realisieren gemeinsam eine homogene Verhaltenskontrolle des gesamten Systems. Somit haben wir einen ersten Schritt bei der Erforschung hierarchischer inkrementeller und lernender Systeme realisiert und die erste Tragfähigkeit des gewählten Ansatzes gezeigt.

3 Erwartungsbasierte Kontrolle und auditorisches Lernen: ALIS 2

Mit ALIS 1 sind die Grundlagen für die weitere Verfolgung unserer Forschungsstrategie gelegt. Das darauf folgende und aufbauende ALIS 2 realisiert weitere wichtige Schritte in die im weiteren beschriebenen Richtungen. Insbesondere betrachten wir hier die Grundlagen zum Erlernen von Semantik und das Aufbrechen der reaktiven Kontrolle. Die abstrakte Architektur von ALIS 2 ist in Bild 2 dargestellt. Die ersten vier Schichten von ALIS 1 entsprechen den ersten drei Schichten von ALIS 2, wobei nur die Form der Darstellung komprimiert wurde. Damit ist dieselbe Form der Interaktion mittels Salienz und Proto-Objekten realisiert. Gegenüber der visuellen Klassifikation mittels der Erscheinung wie in ALIS 1 werden hier jetzt direkt Eigenschaften der Proto-Objekt-Repräsentation als visuelle Merkmale verwendet. Diese werden weiterhin in Schicht 4 erzeugt. Die neue Schicht 5 realisiert eine in Interaktion lernende Worterkennung. Die Verhaltensgenerierung für den Körper liegt hier nun in Schicht 4, implementiert aber die gleichen Funktionalitäten wie bei ALIS 1. In Schicht 6 wird eine homogene Beschreibung der Merkmale für alle möglichen Kanäle gewählt, also für Worte, verschiedene aus dem Sehsystem gewonnene Merkmale wie „ist bewegt“ oder „relative Raumposition“, aber auch der Aktivitätszustand der Verhalten des Systems. Damit steht eine permanent und autonom generierte homogene Beschreibung der Welt und des eigenen Zustands des Systems zur Verfügung. Auf dieser Grundlage kann die Semantik, also die Bedeutung von Wörtern, gelernt und mentale Konzepte gebildet werden. Dies geschieht in Schicht 7, in der die entsprechenden Assoziationen repräsentiert und aktiviert werden. Bei ALIS 2 haben wir uns zunächst auf das Erlernen der Wörter beschränkt, bei ALIS 3 wird diese Beschränkung aufgelöst werden.

Wie gestaltet sich nun die Interaktion mit diesem System? Der Aufbau der Interaktion erfolgt wie bei ALIS 1 mittels Salienz und Peripersonal-Raum, die weitere Kommunikation erfolgt aber entsprechend den neuen Fähigkeiten. Es können sowohl Gegenstände im Peripersonal-Raum als auch Bewegungen oder horizontale Flächen im weiteren Raum um den Roboter herum fixiert werden. Es gibt eine Auswahl von Eigenschaften, die mit beliebigen Äußerungen als Namen verknüpft werden können. Die Namen werden in Interaktion ohne nennenswerte Zeitverzögerung gelernt. Welche Eigen-

schaft benannt werden soll wird vom Interaktionspartner festgelegt, zur Auswahl stehen die relative Position, der Bewegungszustand, die absolute Höhe einer horizontalen Fläche sowie der Aktionszustand des Roboters. Für die jeweiligen Eigenschaften gibt es verschiedene, hier noch festgelegte Alternativen oder Klassen. Da die Eigenschaften direkt aus den Proto-Objekten heraus generiert werden sind sie unabhängig von der detaillierten Erscheinung. So kann man vor den Roboter treten, ihm einen Gegenstand auf seiner linken Seite zeigen und ihn belehren, diese relative Position mit dem Namen „links“ oder jeder beliebigen anderen Äußerung zu verknüpfen. Es ist möglich verschiedene Äußerungen mit ein und derselben Eigenschaft zu verknüpfen, d.h. es können Synonyme gebildet werden. Nach der Belehrung ist jetzt zum Beispiel die Position links mit einem Namen verknüpft. Diese Verknüpfung kann nun wie folgt zur erwartungsbasierten Kontrolle des Roboters eingesetzt werden. Wenn der Interaktionspartner dem Roboter einen zuvor gelernten Namen nennt, wird über die Assoziation eine Erwartungshaltung, hier z. B. für die Position links vor dem Roboter, aktiviert. Wird diese Erwartungshaltung durch eine entsprechende sensorische Perzeption befriedigt, passt die Erwartungshaltung zum Perzept und ein sogenanntes Match-Verhalten wird ausgelöst, hier bei ALIS 2 ein bestätigendes Nicken des Roboters und Löschen der Erwartung. Passt das Perzept nicht zur Erwartungshaltung des Roboters, wird ein sogenanntes Konfliktauflösungsverhalten ausgelöst. Hier signalisiert der Roboter durch Kopfschütteln und Beenden der Fixation, dass seine Erwartung nicht erfüllt ist. Die Erwartung bleibt aber bis zu einer späteren Erfüllung in der Interaktion bestehen. Dieser Mechanismus wird homogen auf alle externen Perzepte angewandt. Nur bei den internen Perzeptionen bzgl. des Aktionszustandes werden nach der Überprüfung zwischen Perzeption und Erwartung andere Aktionen benutzt. Bei einem Match wird das gerade aktive Verhalten weiter ausgeführt, ansonsten wird das zum perzipierten Namen assoziierte Verhalten aktiviert. Mit andern Worten werden die erlernten Namen direkt als Kommandos eingesetzt.

Der hier verfolgte Ansatz ist am ehesten mit dem von Balkenius zu vergleichen, wobei wir bei der Auswertung der Erwartungen über seinen Vorschlag hinaus gehen [8]. Mehr Details zu ALIS 2 finden sich in [9–11].

Zusammenfassend haben wir damit ein System, welches zunächst autonom und reaktiv agiert und damit stabile Interaktionen mit seiner Umwelt herstellt. Auf dieser Basis werden die Grundlagen für das Erlernen von Semantik gelegt, nämlich die homogene Repräsentation der Perzeptionen und Aktionen sowie die explizite Repräsentation der Assoziationen. Gelernt werden hier zunächst nur die auditorischen Bezeichner, was aber im folgenden System direkt aufgebrochen wird. Oberhalb der reinen Reaktivität wurde eine erwartungsbasierte Kontrolle eingeführt, die die Wahl der Aktionen des Systems weiter von der momentanen Perzeption entkoppeln. Da-

mit sind die Erwartungen ein Schritt hin zum Erwerb von Zielen, die in Termini des Systems definiert sind. Die momentane Implementierung der Konfliktauflösung mittels Gesten ist relativ basal, zeigt aber die Möglichkeiten des Ansatzes auf und kann leicht durch deutlich reichhaltigere und zielgerichtete Verhalten wie z.B. Suchen erweitert werden.

4 Erlernen mentaler Konzepte: ALIS 3

Mit ALIS 3 betrachten wir das momentan letzte der erforschten und implementierten Systeme. In ALIS 2 wurden die Grundlagen für das Erlernen von Semantik gelegt, hier gehen wir weitere Schritte in diese Richtung. Die Architektur von ALIS 3 entspricht der von ALIS 2. Die Erweiterungen liegen hier innerhalb der verschiedenen Schichten, wobei die folgenden besonders hervorgehoben werden sollen. Die Worterkennung wurde mit Hilfe einer gekoppelten audio-visuellen Aufmerksamkeit dahin gehend erweitert, dass sie direkt mit den Ohren des Roboters erfolgen kann und nicht mehr mit einer Sprechgarnitur am Interaktionspartner erfolgen muss. Damit ist ein viel natürlicherer Umgang mit dem Roboter möglich. Konnte in ALIS 2 nur eine neue auditorische Bezeichnung gelernt werden, ist jetzt das gleichzeitige Erlernen der Repräsentationen für Sprache, visuelle Merkmale und deren Assoziationen möglich. Damit sind wir dem Ziel des Erlernens der Semantik und mentaler Konzepte ein weiteres Stück näher gekommen. Erreicht wird dies durch verteilte lokale Steuerung des Lernens und einer losen Kopplung über die Assoziationen. Dabei sind mehrere überlappende hybride Repräsentationen in den verschiedenen Merkmalskanälen mit orthogonalen Gruppen möglich. Konkret bedeutet dies, dass nach dem Hinweis bzgl. des Fokus auf z.B. die relative Position sowohl die visuelle Klasse und der auditorische Name als auch deren Assoziation in der Interaktion gelernt werden. Dabei ist ein Überlapp zwischen den Repräsentationen für z.B. „links“ und „unten“ möglich, wobei nicht konkurrierende orthogonale Gruppen für „links“ und „rechts“ sowie „oben“ und „unten“ autonom aufgebaut werden können. Mit Hilfe dieser Gruppen können bei der Evaluation globale Verallgemeinerungen geschaffen werden. In der Interaktion bedeutet dies, dass über die Trainingsdaten hinaus klassifiziert werden kann ohne auf die notwendige lokale Repräsentation der Lernhistorie für die Erkennung von neuen Daten verzichten zu müssen. Weiterhin wurde das Erkennungsspektrum um die Größe von Proto-Objekten erweitert, damit kann unabhängig von der Erscheinung von Objekten deren Größe klassifiziert werden. Das Verhaltensspektrum wurde um das reflexive Greifen von momentan noch spezifizierten Objekten ergänzt. Das Greifen enthält Komponenten der Planung, Selektion und Optimierung von Trajektorien sowie Griffpositionen [12]. Dies repräsentiert einen Schritt in Richtung physikalischer Interaktion im Sinne von Manipulation und damit

die Möglichkeit erlernte Fähigkeiten anwenden zu können.

Die Interaktion mit ALIS 3 gestaltet sich prinzipiell wie mit ALIS 2 inklusive der erwartungsbasierten Verhaltenskontrolle. Das Erlernen von Semantik wurde teilweise auch schon in den Arbeiten von Roy [13] und Iwahashi [14] adressiert. Mit ALIS 3 ist es zum ersten Mal möglich in freier Interaktion mehrere Repräsentationen gleichzeitig zu lernen und direkt zu evaluieren, und dies unter den Randbedingungen einer sich autonom bewegend Plattform. Eine erste Publikation zu ALIS 3 findet sich in [15].

Details zu den internen Mechanismen von ALIS 3 stellen sich wie folgt dar. Aus den grundlegenden perzeptiven Elementen werden Merkmale f_i^t aus der Domäne i zur Zeit t gewonnen. Eine Domäne i ist momentan eine festgelegt basale Perzeption. Hier betrachten wir die folgenden: Für die Worterkennung ($i=1$) ist es eine Kombination von RASTA-PLP und HIST Merkmalen. Weiterhin gibt es die visuell determinierte 3D Position ($i=2$), die approximierte Größe ($i=3$), den Bewegungstatus ($i=4$) und einige Eigenschaften planarer Objekte ($i=5$) eines fixierten Proto-Objektes sowie den Aktivierungsstatus der basalen reaktiven Verhalten ($i=6$). Die Verhalten bilden jeweils eine geschlossene Visuo-Motor-Schleife. Sie werden arbitriert und können durch einen externen Bias auch ohne den korrespondierenden Stimulus aktiviert werden. Die Verhalten umfassen die Annäherung (approach), die Fixation von und das Zeigen auf ein selektiertes Proto-Objekt (fixate, point), die Rückkehr zur Startposition, das Greifen und die Freigabe von Objekten, Kopfnicken (nod) und Kopfschütteln (shake). Dabei stellt das Greifen das komplexeste Verhalten dar. Die Hände des Roboters sind auf der Innen- und Außenseite mit Entfernungssensoren ausgestattet, die Objekte in direkter Nähe zur Hand detektieren. Bei einer Aktivierung des Handinnensensors erfolgt ein reflexartiges Schließen der Finger, ähnlich zu dem kleiner Kinder. Eine Aktivierung des Sensors an der Handaußenseite initiiert ein Öffnen der Finger. Das Greifen besitzt einen komplexeren Modus, der die Planung und Ausführung von Ganzkörperbewegungen zur Annäherung und zum Greifen einfacher Objekte beinhaltet. Alle Merkmalsvektoren und die Bias der Verhalten bilden das Interface für das Erlernen von Konzepten. Dies beinhaltet das Erlernen von Klassen der einzelnen Merkmale sowie das Erlernen von Assoziationen zwischen den einzelnen Klassen. Mit Hilfe der Assoziationen können Erwartungshaltungen generiert und evaluiert werden.

Im Detail führt die kontinuierliche Perzeption und Merkmalsberechnung der unteren Schichten zu Aktivierungen der Speicher $m_{i,x}^t(T_i)$ der Klasse x der Domäne i basierend auf den Beobachtungen der Zeitspanne T_i . Diese Speicherung erfolgt mit Hilfe lokaler Modelle, was einer genaueren Modellierung der tatsächlich beobachteten Daten zuträglich ist. Für die Generalisierung führen

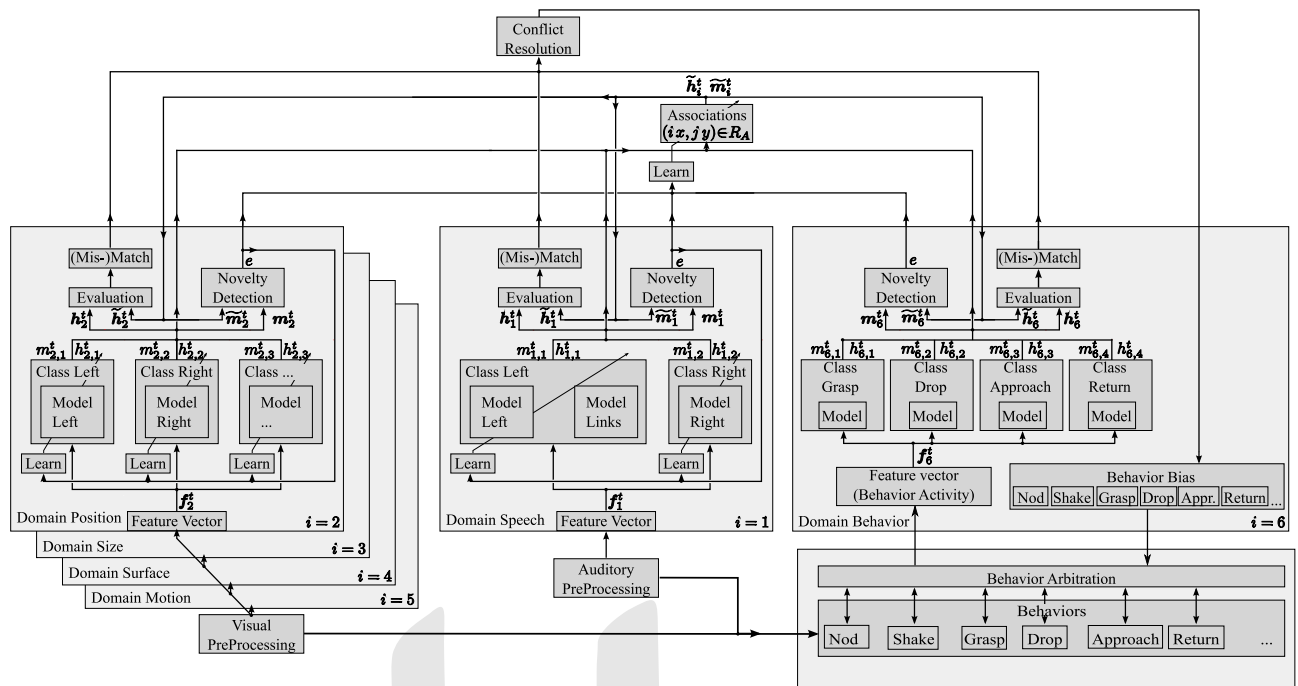


Bild 4 Schematische Darstellung der Mechanismen zum Erlernen mentaler Konzepte.

wir die hypothetische Zuordnung $h_{i,x}$ der Merkmale zu Klassen ein. Diese erhalten wir durch eine vollständige Tessellierung der Merkmale aller Klassen einer Domäne:

$$h_{i,x} = \begin{cases} 1 & \text{falls } x = \operatorname{argmax}_{y \in D_i} (m_{i,y}^t(T_i)) \\ 0 & \text{sonst} \end{cases} \quad (1)$$

Die Aktivierung der Speicher wird zum Lernen neuer Klassen und Assoziationen genutzt, die hypothetische Zuordnung für die Evaluierung und Erwartungsbildung. Die Repräsentation der Informationen ist homogen für alle Domänen, welches eine Grundvoraussetzung für das Bilden von Assoziationen darstellt. Assoziationen werden als Vierer-Tupel (i, x, j, y) repräsentiert.

Neue Klassen werden innerhalb von Lernsessions akquiriert. Für solche eine Session wird mit Hilfe einer Maske a_i festgelegt, welche Domänen am Lernen beteiligt werden sollen. Innerhalb des Zeitrahmens $[t_0, t_1]$ einer Session werden stabile Zuordnungen von Merkmalen zu Klassen angenommen. Die Entscheidung ob eine neue Klasse angelegt oder eine bestehende adaptiert werden soll wird auf Grund des Neuigkeitswertes entschieden. Dabei werden sowohl die momentanen Aktivierungen $m_{i,x}^t(T_i)$ der eigenen Domäne als auch die über die Assoziationen kommunizierten Aktivierungen der anderen Domänen berücksichtigt. Die letzteren nennen wir Top-Down-Aktivierungen $\tilde{m}_{i,x}^t$

$$\tilde{m}_{i,x}^t = \sum_{(j,y) | (i,x,j,y) \in R_A} w_j \cdot m_{j,y}^t(T_j), \quad (2)$$

mit $w_j \in [0, 1]$ als a-priori Gewichtung des Einfluss der Domäne j . Das Gesamtergebnis ist ein Lernsignal e , das

die zu adaptierende Klasse oder die Kreation einer neuen anzeigt:

$$e = \begin{cases} \operatorname{argmax}_{x \in D_i} \tilde{m}_{i,x}^t & \text{falls } \max_{x \in D_i} (\tilde{m}_{i,x}^t) > \Theta_K \\ N_i + 1 & \text{sonst} \end{cases}$$

Dabei wird die Anzahl der existieren Klassen mit N_i bezeichnet, Θ_K bezeichnet einen Schwellwert.

Das Lernen von Assoziationen beruht auf der Gleichzeitigkeit von Beobachtungen. Dabei wird das Ergebnis der Neuigkeitsbewertungen aller in der Lernmaske aktivierten Domänen herangezogen. Um das Lernen zwischen mehr als zwei Domänen zu ermöglichen werden alle Tupel (i, e) für alle aktiven Domänen i mit $a_i > 0$ in der Menge L gespeichert. Die Kombination aller Tupel aus L wird dann zur Menge aller Assoziationen R_A hinzugefügt:

$$R_A = R_A \cup \{(i, x, j, y) | (i, x) \in L, (j, y) \in L\} \quad (3)$$

Dabei sind alle Kombinationen zwischen bestehenden und neuen Klassen möglich.

Die Evaluierung der erlernten Klassen und Assoziationen erfolgt mit Hilfe der Hypothesen, die einen homogenen Vergleich zwischen verschiedenen Domänen erlauben. Dazu werden zunächst alle Hypothesen $h_{i,x}$ für alle Domänen gemäß (1) berechnet. In einem nächsten Schritt wird die Aktivierung einer Klasse mit den Aktivierungen aller assoziierten Klassen verglichen. Dies sollte eine Übereinstimmung ergeben wenn alle assoziierten Klassen gleichzeitig aktiv sind. Im Gegensatz dazu ergibt sich ein Konflikt, wenn eine assoziierte Klasse nicht aktiv

ist und es in der betreffenden Domäne aber eine andere aktive Klasse gibt.

Die mathematische Formulierung dieses Vorgangs erfolgt mit Hilfe der Ähnlichkeit zweier Aktivierungen. Der erste Vektor $\mathbf{h}_i$ ist die Konkatinierung der Hypothesen in Domäne i :

$$\mathbf{h}_i = (h_{i,1}, \dots, h_{i,N_i})^T. \quad (4)$$

Der zweite Vektor repräsentiert die Top-Down Hypothese $\tilde{\mathbf{h}}_i$:

$$\tilde{\mathbf{h}}_i = (\tilde{h}_{i,1}, \dots, \tilde{h}_{i,N_i})^T, \quad (5)$$

mit

$$\tilde{h}_{i,x} = \sum_{\substack{(j,y) | (i,x,j,y) \in R_A \\ \wedge (j,y) \neq (i,x)}} h_{j,y}. \quad (6)$$

Die Top-Down-Hypothese $\tilde{\mathbf{h}}_{i,x}$ ist die Summation aller mit $h_{i,x}$ assoziierten Hypothesen, jedoch ohne $h_{i,x}$ selbst. Der Vergleich kann mit Hilfe des Skalarprodukts $d(\mathbf{h}_i, \tilde{\mathbf{h}}_i)_S$ mit nachfolgender Schwellwertbildung erfolgen. Der so berechnete Vergleich kann direkt zur Evaluierung der erlernten Klassen und Assoziationen genutzt werden. Hierbei unterscheiden wir momentan zwischen einer perzeptiven und einer verhaltenskontrollierenden Domäne.

Eine perzeptuelle Übereinstimmung/Konflikt erfolgt wenn ein Tutor dem System ein Objekt mit einer bestimmten visuellen Klasse zeigt und eine entsprechende Bezeichnung ausspricht. Im Falle der Übereinstimmung wird ein Bias für das Nicken des Roboterkopfes ausgelöst und die Evaluierung ist beendet. Im Falle eines Konfliktes wird ein Bias für das Kopfschütteln ausgelöst und ein anderes Objekt fixiert, wobei die Erwartung ein Objekt mit einer dem Bezeichner entsprechenden Eigenschaft zu sehen weiterhin bestehen bleibt. Das Auslösen der Kopfbewegungen ist jeweils nur ein Beispiel für eine allgemeinere Konfliktlösung. Noch nicht realisierte Alternativen sind das Suchen nach Objekten mit bestimmten Eigenschaften oder das Fragen nach solchen. Die persistente Aktivierung mittels der Assoziationen stellt aber bereits einen einfachen Mechanismus zur Realisierung von Zielen dar.

Eine Evaluierung mit einer verhaltenskontrollierenden Domäne ergibt sich, wenn ein Tutor einen Bezeichner nennt, der mit einem Verhalten assoziiert ist. Falls dieses Verhalten bereits aktiv ist liegt eine Übereinstimmung vor und die Evaluation ist beendet. Im Falle eines Konfliktes wird ein Bias für das assoziierte Verhalten ausgelöst. Dadurch hat man in der homogenen Architektur eine natürliche Möglichkeit zur Realisierung von Kommandos.

Was sind die besonderen Eigenschaften dieses Ansatzes? Herauszuheben sind die homogenen Repräsentationen der Klassen und Aktivierungen, die das lokale Lernen und die Top-Down gestützte Hypothesenbildung erlauben. Die Architektur realisiert das gleichzeitige Lernen in verschiedenen Domänen und deren Assoziation während

der Interaktion mit dem Roboter. Der Ansatz ist skalierbar und offen für weitere Top-Down Aktivierungen von zukünftigen Schichten.

5 Erweitertes Spektrum und Ausblick

Die bisher beschriebenen Forschungsaktivitäten werden durch weitere exploratorische Arbeiten ergänzt und erweitert. Bisher wurde z. B. das Erlernen von Verhalten noch nicht entsprechend adressiert. Einen Ansatzpunkt bietet hier das Lernen durch Imitation. Erlernte Verhalten können dann dem Verhaltensspektrum hinzugefügt und autonom weiter verfeinert werden. Erste Arbeiten dazu sind in [16; 17] dargestellt. Ein anderer Ansatzpunkt ist das Lernen durch direkte Unterweisung mit anschließendem Extrahieren der unbedingten Regelmäßigkeiten [18].

6 Diskussion und Zusammenfassung

In diesem Artikel haben wir eine Übersicht über unsere Forschungsarbeiten auf dem Gebiet der kognitiven humanoiden Robotik gegeben. Warum haben wir den oben beschriebenen Weg für unsere Forschung gewählt? Wir sind der Meinung, dass Fortschritte in Bezug auf Robustheit und Flexibilität wie beim biologischen Vorbild nur möglich sind, wenn ein entsprechend inkrementeller und hierarchischer Ansatz verfolgt wird. Dies korrespondiert teilweise mit dem Ansatz der Entwicklungsrobotik [19; 20], wobei wir hier nicht alle Entwicklungsstufen eines Artefakts vollständig autonom durchlaufen wissen wollen, sondern ein entsprechendes wissenschaftliches Verständnis der Zwischenstufen und der zu Grunde liegenden architekturellen Prinzipien anstreben. Dabei sind wir überzeugt, dass wir unsere Ansätze bereits zu einem frühen Zeitpunkt mit der realen Welt und Interaktionen mit realen Menschen konfrontieren müssen. Die Forschungsökonomie gebietet natürlich den Einsatz von Simulationswerkzeugen, aber nur ein früher Schritt in die Zielumgebung gewährleistet ein hinreichend starkes Korrektiv für die gestellten Forschungsfragen und gewählten Ansätze. Dieser Übergang von notwendigen akademischen Modellbildungen hin zu belastbaren Beiträgen in der realen Welt ist auch als Skalierungsproblematik bekannt. Ohne sich diesem Problem zu stellen wird es keinen Fortschritt in Richtung der Eingangs genannten Ziele geben.

Danksagung

Die Autoren möchten sich bei den Kollegen des Honda Research Institute Europe für ihre motivierte Forschungsarbeit, über die hier berichtet wird, bedanken.

Literatur

- [1] O. Brock, A. Fagg, R. Grun, R. Platt, M. Rosenstein, and J. Sweetney. A framework for learning and control in intelligent humanoid robots. *International Journal of Humanoid Robotics*, 2(3), 2005.

- [2] R. C. Arkin, M. Fujita, T. Takagi, and R. Hasegawa. An ethological and emotional basis for human-robot interaction. *Robotics and Autonomous Systems*, (3–4):191–201, 2003.
- [3] S. Chernova and R. C. Arkin. From deliberative to routine behaviors: A cognitively inspired action-selection mechanism for routine behavior capture. *Adaptive Behavior*, 15(2):199–216, 2007.
- [4] T. Asfour, K. Regenstein, P. Azad, J. Schröder, A. Bierbaum, N. Vahrenkamp, and R. Dillmann. ARMAR-III: An integrated humanoid platform for sensory-motor control. In *Proceedings of the IEEE/RSJ International Conference on Humanoid Robots (Humanoids 2006)*, Genoa, Italy, 2006.
- [5] D. Vernon, G. Metta, and G. Sandini. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *IEEE Transactions on Evolutionary Computation*, 11(2):151–180, 2007.
- [6] C. Goerick, B. Bolder, H. Janssen, M. Gienger, H. Sugiura, M. Dunn, I. Mikhailova, T. Rodemann, H. Wersing, and S. Kirstein. Towards incremental hierarchical behavior generation for humanoids. In *IEEE-RAS International Conference on Humanoids 2007*. IEEE, 2007.
- [7] A. Ude and G. Cheng. Object recognition on humanoids with foveated vision. In *Proceedings of the IEEE/RSJ International Conference on Humanoid Robots (Humanoids 2004)*, Los Angeles, 2004.
- [8] C. Balkenius. *Natural Intelligence in Artificial Creatures*. Lund University Cognitive Studies 37, 1995.
- [9] I. Mikhailova, M. Heracles, B. Bolder, H. Janssen, H. Brandl, J. Schmüdderich, and C. Goerick. Coupling of mental concepts to a reactive layer: incremental approach in system design. In *Proceedings of the 8th International Workshop on Epigenetic Robotics, Brighton, UK*, 2008.
- [10] B. Bolder, H. Brandl, M. Heracles, H. Janssen, I. Mikhailova, J. Schmüdderich, and C. Goerick. Expectation-driven autonomous learning and interaction system. In *IEEE-RAS International Conference on Humanoids*, 2008.
- [11] J. Schmüdderich, H. Brandl, B. Bolder, M. Heracles, H. Janssen, I. Mikhailova, and C. Goerick. Organizing multimodal perception for autonomous learning and interactive systems. In *IEEE-RAS International Conference on Humanoid Robots*, 2008.
- [12] M. Gienger, M. Toussaint, N. Jetchev, A. Bendig, and C. Goerick. Optimization of fluent approach and grasp motions. In *International Conference on Humanoid Robots*. IEEE Press, 2008.
- [13] D. Roy. *A mechanistic model of three facets of meaning*. Oxford University Press, 2008.
- [14] N. Iwahashi. Robots that learn language: Developmental approach to human-machine conversations. In P. Vogt et al., editors, *Symbol Grounding and Beyond: Proceedings of the Third International Workshop on the Emergence and Evolution of Linguistic Communication*, pages 143–167. Springer, 2006.
- [15] C. Goerick, J. Schmuedderich, B. Bolder, H. Janssen, M. Gienger, A. Bendig, M. Heckmann, T. Rodemann, H. Brandl, X. Domont, and I. Mikhailova. Interactive online multimodal association for internal concept building in humanoids. In *IEEE-RAS International Conference on Humanoids 2009*. IEEE, 2009.
- [16] M. Muehlig, M. Gienger, S. Hellbach, J. Steil, and C. Goerick. Task-level imitation learning using variance-based movement optimization. In *International Conference on Robotics and Automation (ICRA)*. IEEE, 2009.
- [17] B. Dariush, M. Gienger, A. Arumbakkam, Y. Zhu, B. Jian, K. Fujimura, and C. Goerick. Online transfer of human motion to humanoids. *International Journal of Humanoid Robotics*, 2009.
- [18] M. Howard, S. Klanke, M. Gienger, C. Goerick, and S. Vijayakumar. Behaviour generation in humanoids by learning potential-based policies from constrained motion. *Applied Bionics and Biomechanics*, – Special Issue on Humanoid Robots, 2009.
- [19] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291(5504):599–600, 2000.
- [20] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini. Developmental robotics: A survey. *Connection Science*, 15(4):151–190, 2003.

Manuskripteingang: 30. September 2009



Dr.-Ing. Christian Goerick ist Chief Scientist am Honda Research Institute Europe. Hauptarbeitsgebiete: Kognitive Robotik, Fahrerassistenzsysteme, Systemarchitekturen, eingebettete Seh- und Verhaltenssysteme.

Adresse: Honda Research Institute Europe, Carl-Legien-Straße 30, D-34109 Offenbach/Main, Fax: +49-(0)69-89011-759,
E-Mail: christian.goerick@honda-ri.de



Prof. Dr.-Ing. habil. Edgar Körner ist Präsident des Honda Research Institute Europe. Hauptarbeitsgebiete: Brain-like Intelligence, Selbst-referentielle Kontrollarchitekturen, Selbst-organisierende Wissensrepräsentationen, Kognitive Robotik.

Adresse: Honda Research Institute Europe, Carl-Legien-Straße 30, D-34109 Offenbach/Main, Fax: +49-(0)69-89011-759,
E-Mail: edgar.koerner@honda-ri.de