

Adaptive Intelligent Systems Assistance (AISA)

Marc Rettschlag

rettschlag@ifm-mv.de

IFM Institut für Multimediatechnik gGmbH
Alter Holzhafen 17b
23966 Wismar

Abstract: Die Adaptive Intelligent Systems Assistance (AISA) ist ein Projekt des Instituts für Multimediatechnik (IFM). Ziel dieses Projektes ist es, heutige technologische Möglichkeiten der Informationsbeschaffung, Vernetzung und Steuerung mit dem Kontext eines Anwenders zu koppeln, um so Bedienelemente und Informationen zielgerichtet zu generieren oder Eingaben auszuwerten. Der Kontext eines Anwenders kann sich dabei aus dem Aufenthaltsort, der Auswahl eines interaktiven Systems und aus dem Interaktionsverlauf mit einem System ergeben. Zudem können Gewohnheiten oder das Verhalten in ähnlichen Situationen ausgewertet und verwendet werden. Mit der Anwendung eines solchen Konzeptes auf mobilen Endgeräten und mit dem Einsatz von Spracherkennung und Sprachsynthese soll eine schnellere, einfachere und benutzerzentrierte Mensch-Maschine Interaktion erreicht werden.

1 Einleitung

Die fortschreitende Technisierung und Vernetzung der Umwelt schafft viele Möglichkeiten für den Menschen. Neben Handys und Computern gibt es mittlerweile intelligente Kühlschränke, programmierbare Heizungs- und Beleuchtungssysteme, die das alltägliche Leben vereinfachen oder bequemer gestalten sollen. Allerdings bringen diese neuen Geräte auch Hindernisse bei der Anwendung mit sich. Mit der Menge an Geräten und Funktionen nimmt auch die Anzahl und Komplexität der Benutzerschnittstellen zu. Fernbedienungen mit dutzenden von Tasten oder Touchpanels mit vielen unübersichtlichen Untermenüs stellen keine komfortable Steuerung bereit.

Ursache dafür ist unter anderem die fehlende Interoperabilität von Geräten und Gerätesystemen der verschiedenen Bereiche (Heimautomation, Heizungssteuerung, Heimelektronik), so dass eine Kopplung und Steuerung über ein gemeinsames Interface nicht möglich ist. Allerdings ist die bedarfsgerechte Gestaltung der Benutzerschnittstelle wesentlich für die Einbindung von Technologien der Informations- und Kommunikationstechnik [HM09], gerade im Bereich *Ambient Assisted Living*. Hier setzt das AISA-Projekt an. Die Benutzerschnittstelle soll sich den Umgebungsgegebenheiten des Anwenders anpas-

sen bzw. bei Eingaben diese auswerten können und damit die Usability¹ erhöhen. Für diese kontextbasierte Adaption werden die für das System relevanten und erfassbaren Charakteristika der aktuellen Situation ausgewertet. Beispielsweise kann sich der Kontext eines Anwenders aus seinem Aufenthaltsort, mit Hilfe von Sensoren aufgenommenen Messwerten und Parametern des zu steuernden Systems ergeben. Zudem besitzt das System eine Art Kurzzeit- und Langzeitgedächtnis, um den vorangegangenen Interaktionsverlauf mit einem System oder die Gewohnheiten und Vorlieben des Nutzers auszuwerten. Die Nutzerinteraktion findet über ein mobiles Endgerät, die Erkennung natürlicher Sprache und die Ausgabe von Sprache statt. Mit Hilfe der Kontextinformationen werden dann im Fall einer grafischen Oberfläche passende Bedienelemente oder Bedienungsabläufe generiert oder bei einer Spracheingabe diese ausgewertet. Im Idealfall werden diese Kommunikationskanäle nicht alleinig, sondern sich ergänzend in Kombination genutzt.

Wichtig beim Einsatz von Spracherkennung ist es, den Benutzer möglichst frei agieren zu lassen. Gerade bei älteren Nutzern sollte vermieden werden, eine Barriere aufzubauen, beispielsweise durch einen vorgefertigten Satz von Kommandos oder eine einzuhaltende Befehlsreihenfolge. Gleichzeitig muss das Systemverhalten nachvollziehbar sein, so dass der Nutzer nicht irritiert oder überrascht wird und sich unsicher ist, welchen Effekt seine Eingaben haben könnten. Zudem ist auch die Frage zu stellen, wie natürlich und „menschlich“ ein Nutzer überhaupt mit einem technischen System interagieren möchte.

Die Umsetzung des Systems erstreckt sich über verschiedene Fachbereiche. So muss die Hardware (Heimelektronik, Beleuchtung, Sensoren, Aktoren) angebunden werden, dies erfordert die Integration heterogener Netzstrukturen (Bussysteme, lokale Netzwerke) und Systeme. Des Weiteren werden Konzepte aus dem Gebiet der künstlichen Intelligenz und der Computerlinguistik angewendet. Der Schwerpunkt in diesem Beitrag liegt bei natürlicher sprachlicher Eingabe und der Abbildung und Auswertung der Kontextinformationen im AISA-Framework.

2 Kontext und Kontextsensitivität

Die Gewinnung und Nutzung des Kontextes ist ein wesentlicher Punkt des AISA-Frameworks. Als *context-awareness* (Kontextsensitivität) wird das Erkennen von Umgebungsveränderungen durch ein Computersystem bezeichnet, dieser Begriff ist eng verknüpft mit dem *ubiquitous computing*². *Context-adaption* (Kontext-Anpassung) beschreibt die Anpassung des Systems in Struktur, Funktionalität oder Verhalten an die Umgebungsgegebenheiten. Der Kontext umfasst dabei alle Informationen, die die Situation einer Entität beschreiben. Eine Entität kann eine Person, ein Objekt oder ein Ort sein, der für eine Interaktion zwischen Nutzer und Anwendung relevant ist. Weiterhin kann zwischen dem direkten und indirekten Kontext unterschieden werden [GPZ05]. Der direkte Kontext ergibt sich aus den wahrnehmbaren Informationen, wie Sensordaten, Position oder Gerätestatus

¹Usability - Gebrauchstauglichkeit eines Produkts in einem Nutzungskontext. Ziele: Effektivität, Effizienz und Zufriedenheit der Nutzer.

²*Ubiquitous computing* bezeichnet das allgegenwärtige Vorhandensein von rechnergestützter Informationsverarbeitung.

und definierten Informationen, wie die übliche Aufstehzeit, Essgewohnheiten oder Tableteneinnahmezeit. Durch Interpretation dieser Informationen kann der indirekte Kontext abgeleitet werden. Als einfaches Beispiel könnten die Informationen PersonX(Ort: Wohnzimmer) und TV(Status: An, Ort: Wohnzimmer) den Schluss zulassen, dass die Person gerade Fernsehen guckt.

Für die computergestützte Gewinnung des Kontextes muss die reale Welt abstrahiert und in eine elektronische Form gebracht werden, so dass eine Anwendung diese Informationen auswerten kann. Hierfür gibt es verschiedene Kontextmodelle, welche die Entitäten, ihre Beziehungen, Eigenschaften und zeitlichen Veränderungen abbilden. Im AISA-Framework wurde ein ontologiebasierter Ansatz gewählt. Gründe dafür sind der verteilte Aufbau, Validierungsmöglichkeiten, die Beschreibungsmöglichkeiten der Interaktionspartner und die gute Anwendbarkeit auf existierende Umgebungen. Andere Möglichkeiten wären zum Beispiel grafische, logikbasierte oder objektorientierte Modelle gewesen [SLP04].

Ontologien stammen aus dem Gebiet der Künstlichen Intelligenz und können als Möglichkeit der Wissensrepräsentation betrachtet werden. Sie dienen der Strukturierung, Speicherung und dem Austausch von Daten in bestehenden Wissensbeständen. Sie beschreiben eine Konzeptualisierung einer Domäne und bestehen aus einer formalen, expliziten Beschreibung von relevanten Begriffen (engl.: *concepts*), sowie ihrer Eigenschaften und Beziehungen untereinander. Im Rahmen des AISA-Frameworks beschränkt sich die abzubildende Welt auf ein Gebäude bzw. einen Gebäudekomplex, mit Etagen, Räumen, Mobiliar, technischen Geräten und weiteren typischen Bestandteilen.

3 Spracherkennung und Sprachkontext

Neben der Interaktion über ein mobiles Endgerät bietet das AISA-Framework auch die Möglichkeit der Kommunikation über natürliche Spracheingaben. In Bild 1 ist der grundsätzliche Ablauf von Spracherkennung und Sprachverstehen zu sehen.

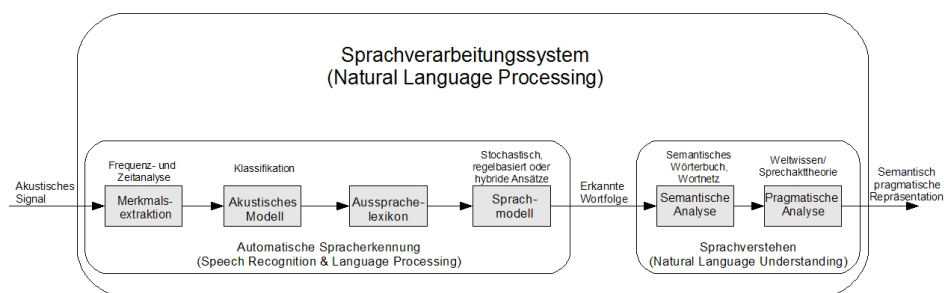


Abbildung 1: Komponenten eines Sprachverarbeitungssystems

Der Prozess der automatischen Spracherkennung beschreibt den Vorgang vom akustischen Signal bis zur erkannten Wortfolge, daran schließt sich der Prozess des Sprachverstehens an. Dieser Vorgang extrahiert neues Wissen aus der Wortfolge. Ziel ist die Bedeutung der natürlich-sprachlichen Äußerung zu finden. Um die Semantik³ zu extrahieren wird Syntaxwissen (Wissen über mögliche auftretende Satzkonstruktionen) und lexikalisches Wissen benötigt. Die pragmatische Analyse bezieht voran gegangene Aussagen mit ein. Durch die semantisch-pragmatische Analyse wird die Bedeutung des eingegebenen Satzes in eine maschinenverwertbare Form gebracht.

Gesprochene Sprache ist im Gegensatz zu geschriebener meist situationsgebunden, der Sprecher richtet sich an einen Gesprächspartner, der Teil der Situation ist und mit dem Wissen geteilt wird. Die reine Wortfolge macht nur einen Teil des dieses Wissens aus und reicht nicht aus, um eine eindeutige Handlung zu erkennen. Neben dem Gesprächskontext muss zur besseren Interpretation der Umgebungskontext ausgewertet werden.

4 Anbindung der Hardware

Der Umgebungskontext ergibt sich unter anderem aus den vorhandenen Geräten und angeschlossenen Sensoren. Im Rahmen des Projektes wurde ein Framework entwickelt, welches über eine flexible Schnittstelle zu bestehenden Infrastrukturen (Bussysteme, Netzwerke) verfügt. Die Prioritäten dieser *Automation Framework* (siehe Abbildung 2) genannten Schnittstelle lagen bei der Verwendung von bestehenden Standardtechnologien und bei einer möglichst breiten Anbindung von verschiedenen Geräteklassen, zum Beispiel auch Geräten der Heim- und Unterhaltungselektronik, die keinen Busanschluss besitzen. Für Geräte der Automatisierungstechnik bot sich der Einsatz von *OPC*⁴ an. *OPC* ist eine standardisierte Schnittstelle für einen herstellerunabhängigen und leistungsfähigen Austausch von Prozessdaten. Damit ist es möglich, Komponenten der Automatisierungstechnik (wie Sensoren, Regler und Steuerungen) von verschiedenen Herstellern gemeinsam in einem Netzwerk zu nutzen, ohne aufwändige Protokollanpassungen durchzuführen. *OPC* definiert verschiedene Spezifikationen, welche mit Hilfe vieler Firmen aus dem Bereich der Automatisierungstechnik entwickelt wurden⁵. Der *OPC*-Server bietet dem Client einen transparenten Zugriff auf die verschiedenen Datenquellen (Sensor) und Datensinken (Aktor). Der Client ist Bestandteil des *Automation Frameworks*, der Server ist eine externe Software.

Ein weiterer unterstützter Standard für die Hardwareanbindung, der eher im Bereich der Heimelektronik liegt, ist *Universal Plug and Play (UPnP)*. Ähnlich wie beim klassischen Plug and Play (z.B. USB-Maus am PC), macht sich das Gerät selbst bekannt und stellt dann seine Funktionalitäten zur Verfügung, allerdings nicht innerhalb eines Rechners sondern innerhalb eines Netzwerkes (LAN, WLAN, Bluetooth). Dadurch bringt UPnP eine

³Die linguistische Semantik befasst sich mit der Bedeutung sprachlicher Zeichen.

⁴*OPC* stand ursprünglich für *OLE for Process Control*, wird heute mit aber mit *Openness, Productivity and Collaboration* oder ohne Verweis auf eine Abkürzung angegeben.

⁵Momentan gehören knapp 450 Firmen zur *OPC Foundation*, die für die Entwicklung und Verbreitung von *OPC* zuständig ist.

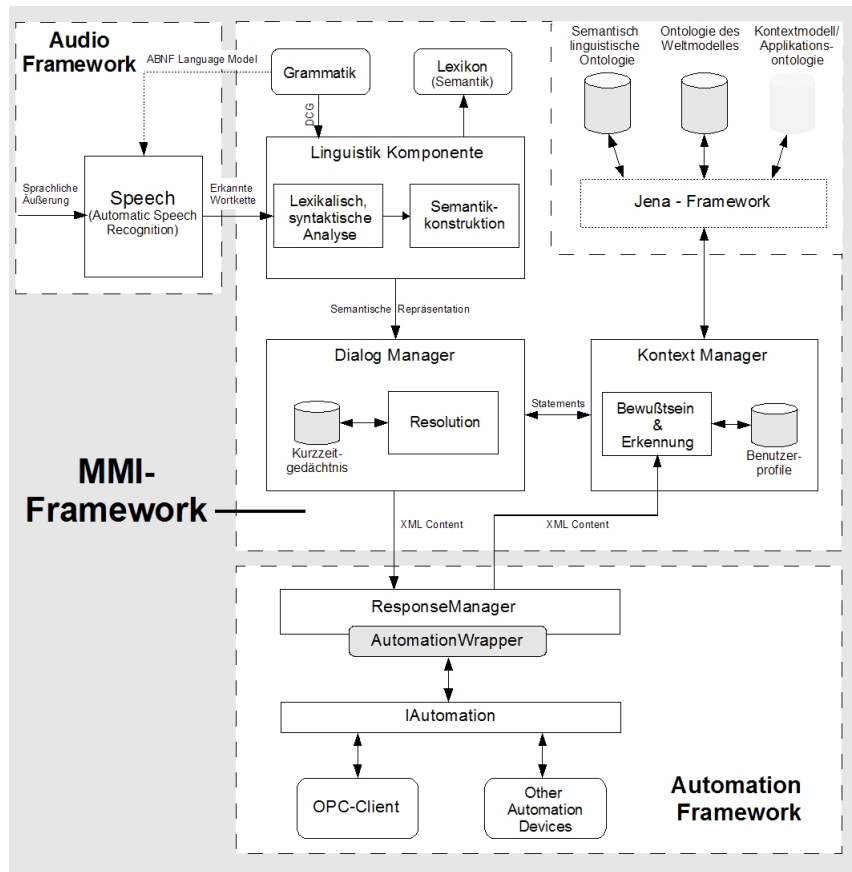


Abbildung 2: Übersicht des AISA-Frameworks

hohe Flexibilität mit und reduziert den Konfigurationsaufwand, da die Einbindung neuer Geräte ohne Eingaben des Nutzers geschehen kann. UPnP nutzt standardisierte Protokolle und Technologien (TCP, UDP, IP, XML, SOAP). Es ermöglicht ähnlich wie OPC eine herstellerübergreifende Ansteuerung von Geräten. Herausgegeben wird der Standard vom UPnP Forum, dem über 800 Firmen aus verschiedenen Bereichen (Computer, Netzwerk, Automation, mobile Geräte) angehören.

Parallel zur Kommunikation über die Standards *OPC* und *UPnP* bietet das Framework noch weitere Mechanismen, um mit Geräten zu interagieren. Dies betrifft vor allem Geräte der Heimelektronik, die nicht über einen Feldbus oder Netzwerkanschluss, aber beispielsweise durch eine RS-232-Schnittstelle steuerbar sind. Dazu wird innerhalb des Frameworks ein an die Schnittstelle angepasster Client implementiert. Der Datenaustausch aller Varianten erfolgt ereignisorientiert, so dass Statusänderungen direkt übertragen werden können und aktuell sind.

Neben der Anbindung über die Clients ist das *Automation-Framework* dafür zuständig

die angebundenen Geräte in ein einheitliches Datenformat zu bringen, so dass innerhalb des Frameworks nicht nach der Art der Anbindung unterschieden wird. Zusätzlich werden weitere Information gespeichert, wie Ort, Eigenschaften und der Gerätestatus. Diese Daten bilden einen Teil des Umgebungskontextes.

5 Kontextauswertung im AISA-Framework

Nachdem in den vorherigen Abschnitten die relevanten Einzelaspekte dargestellt wurden, soll im folgenden etwas genauer gezeigt werden, wie diese im AISA-Framework zusammen wirken. Bild 2 zeigt die grundlegende Struktur des Frameworks.

Im Audio-Framework findet die Spracherkennung statt. Für die Modellierung der möglichen Satzkonstruktionen wird eine Grammatik verwendet. Dies bietet den Vorteil, dass die erkannten Sätze oder Wortfolgen für die weitere Verarbeitung in einer grammatikalisch richtigen Form sind. Nachteilig ist die steigende Komplexität, um möglichst viele Satzkonstruktionen abzubilden. Da der Nutzer vorrangig Anweisungen gibt, liegt der Schwerpunkt bei Imperativsätzen.

In der Linguistik Komponente wird die erkannte Wortkette in Prolog⁶ über eine Definite Clause Grammar (DCG) auf Korrektheit geprüft. Über die Verbindung mit dem Vollformenlexikon⁷ werden die linguistischen Einheiten (z.B. semantische Repräsentation, Wortart) zu den einzelnen Wörtern abgerufen und innerhalb des Satzes in Verbindung gebracht. Eine beispielhafte Eingabe wäre zum Beispiel: „*Schalte das Licht an der Wand aus!*“. Die semantische Repräsentation des Satzes wäre wie folgt:

[op:switch, sem: x], ([quant:one], [device:licht], [sem:location, sem:relativ], [quant:one], [location:wand]), [op:off, sem:todo_X]

Die Parameter definieren sich folgendermaßen:

- Positionsbeschreibung - [location:?]
- Quantor - [quan:?]
- Gerät - [device:?]
- Schaltaktion - [op:?]
- Pronoun - [pron:?]

Die Position wird über das Wort „*Wand*“ bestimmt. Der Quantor gibt an ob ein oder mehrere Geräte geschaltet werden sollen. Das Gerät selbst wird über das Wort „*Licht*“ be-

⁶Prolog ist eine logikbasierte Programmiersprache, die unter anderem in der Computerlinguistik Anwendung findet.

⁷Das Vollformenlexikon enthält weitere semantische Informationen, wie die unterschiedlichen Bedeutungen eines Wortes in Abhängigkeit der Wortstellung innerhalb eines Satzes.

schrieben. Mit dem Wort „aus“ wird die Schaltaktion beschrieben. Da das Wort „an“ ebenfalls im Satz vorkommt und als Schaltaktion interpretiert werden könnte, wird mit Hilfe des Vollformenlexikon die Bedeutung des Wortes in Abhängigkeit von der Position innerhalb des Satzes untersucht. Dadurch wird „an“ als Beziehung (Präposition) interpretiert. Hierbei ist zu erkennen, dass in diesem Fall ein simples *Keyword Spotting*⁸ nicht ausreicht. Durch die Sequenz „Licht an“ würde dies genau den falschen Vorgang auslösen. Die erkannte semantische Repräsentation gibt Aufschluss über den aktuell geäußerten Befehl und wird an den Dialogmanager übergeben. Im Dialogmanager ist ein Kurzzeitgedächtnis implementiert, welches die vorherigen semantischen Repräsentationen mit einem Zeitstempel speichert. Lässt sich eine Äußerung nicht vollständig in eine Befehlsrepräsentation umwandeln, greifen die tiefensemantische Auflösung und die Ellipsenauflösung⁹ darauf zurück. Zudem können über diesen Weg Korrekturen, Ergänzungen oder Verknüpfungen aufgelöst werden, wie sie bei Spontansprache häufig entstehen (z.B. „Mach hier das Licht aus!“ „In der Küche auch!“). Das Zeitfenster kann dabei statisch sein oder vom Gerät oder Nutzerverhalten abhängen.

Für eine Befehlsausführung fehlen allerdings weiterhin Informationen. Zwar sind Gerät („Licht“) und Ort („Wand“) genannt worden, für ein technisches System, welches ein ganzes Haus verwaltet ist das jedoch zu wenig. Die fehlenden Informationen werden über den Kontextmanager und die Ontologie gefunden.

Der Kontextmanager ist ein zentraler Punkt im AISA-Framework, hier werden die Teilspekte des Kontextes (Gesprächskontext und Umgebungskontext) zusammengeführt und mit der realen Welt über das Automationframework verbunden. Die verwendete Ontologie besteht aus drei Modellen, dem Weltmodell, dem semantisch-linguistischen Modell und dem Kontextmodell innerhalb des Interaktionsmanagements¹⁰. Angebunden ist die Ontologie über das Jena-Framework¹¹. Dies ist eine Java API für Manipulationen und Anfragen auf Ontologien. Es bietet verschiedene *Reasoner* für die Ableitung von Schlussfolgerungen und Wissen.

Ontologie des Weltmodells

Das Weltmodell besteht, wie schon erwähnt, aus der Abbildung eines Gebäudes mit Etagen, Räumen, Möbeln, technischen Geräten und Personen. Bild 3 zeigt einige Entitäten und ihre Relationen innerhalb des Modells. Über die Relationen lassen sich logische Verbindungen zwischen den Begriffen finden. So können sich beispielsweise Personen und Gegenstände in einem Bereich(Area) aufhalten oder in der Nähe(isnearto) von anderen Geräten aufhalten.

Das Weltmodell beinhaltet immer die aktuellen Werte(Status, Position) der Geräte und des Nutzers und ist über den Kontextmanager direkt mit Automationframework und da-

⁸Beim Keyword Spotting wird auf eine Grammatik verzichtet und nur das erkannte Wort oder eine bestimmte Wortfolge (Kommando-Sequenz) weitergegeben.

⁹In der Linguistik bezeichnen Ellipsen das Auslassen von Satzteilen, die inhaltlich entbehrlich sind.

¹⁰Das Kontextmodell des Interaktionsmanagements ist hauptsächlich für die grafische Benutzerschnittstelle relevant und wird daher hier nicht näher erläutert

¹¹Jena - A Semantic Web Framework for Java: <http://jena.sourceforge.net>

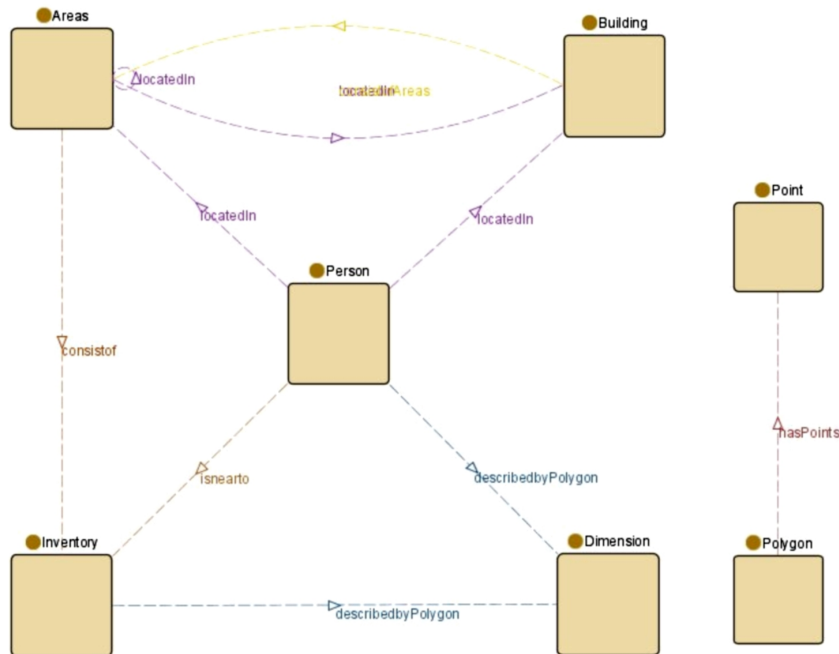


Abbildung 3: Entitäten und ihre Beziehungen im Weltmodell im AISA-Framework

durch mit den echten, physikalisch vorhandenen Geräten verbunden. Das in der Ontologie vorhandene Weltmodell ist somit synchron mit der realen Welt.

semantisch linguistische Ontologie

Die Wordnet genannte semantisch-linguistische Ontologie (Abbildung 4) beinhaltet semantische Beschreibungen für Bereiche, Geräte und Personen. Durch Abbildung des Weltmodells und der Wort-Semantik-Ebene innerhalb einer Ontologie lassen sich beide Modelle effizient verbinden. Zudem ist eine modellbasierte und komfortable Erzeugung von Relationen zwischen Objekten beider Modelle möglich. Das Wordnet mit allen Unterkonzepten und Individuen lässt sich somit auch problemlos mit dem Weltmodell und dessen Objekten in Beziehung bringen. Im konkreten Beispiel bedeutet dies, dass das Wordnet erkennt, was mit „Licht“ gemeint ist. Das Weltmodell beinhaltet zudem die Position und den Status aller Lampen, ein Modell der Räume und die Position des Benutzers. Durch logische Schlussfolgerung (*Reasoner*) kann das System die fehlenden Informationen ergänzen und über das Automationframework die entsprechenden Geräte schalten.

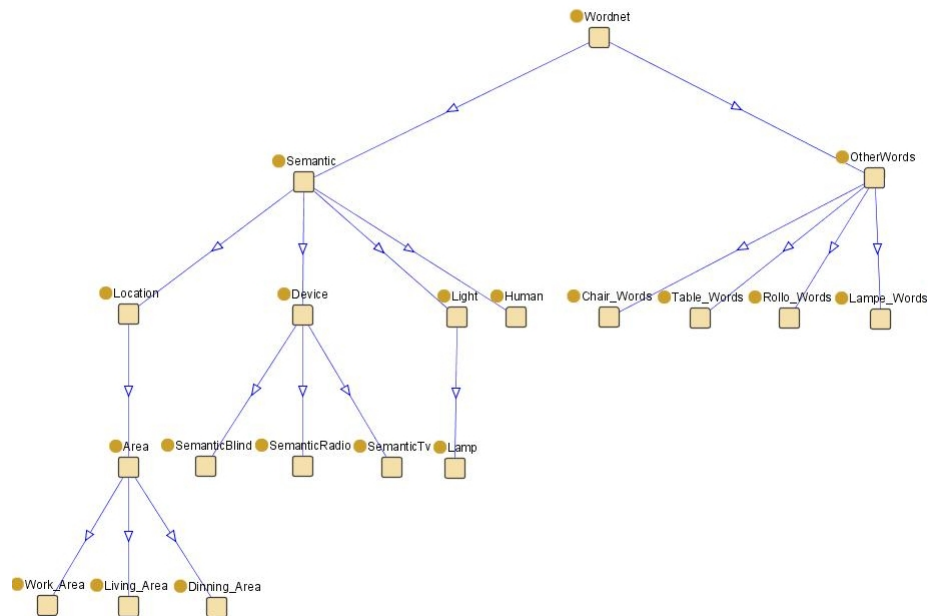


Abbildung 4: semantisch-linguistische Ontologie des AISA-Frameworks

6 Zusammenfassung

Mit dem hier vorgestellten Ansatz der Verbindung des sprachlichen Kontextes mit dem situativen lässt sich eine intuitivere, sprachliche Benutzerschnittstelle realisieren. Da sich das Projekt noch in der Entwicklung findet und daher noch keine ausreichenden, langfristigen Feldversuche in Wohnumgebungen gemacht werden konnten, kann keine abschließende Aussage über die Benutzerakzeptanz gemacht werden. Allerdings wurde, um möglichst früh und fortlaufend Erfahrungen zu sammeln, ein Demonstrator entwickelt, der einige typische Bestandteile einer häuslichen Umgebung enthält (Lampen, elektrische Jalousie), die über ein Bussystem (EIB/KNX) an das AISA-Framework angeschlossen sind. Dieser Demonstrator wurde intern und extern (Messen, Tag der Barrierefreiheit in Berlin) von vielen Personen getestet. Die Personen wurden nicht eigewiesen, sondern konnten frei sprechen. Dabei konnten äußerst lange und förmliche Aussagen („Schalten Sie bitte das Licht ein“) genauso beobachtet werden wie sehr kurze („An“). In vielen Fällen schien den Testern intuitiv klar zu sein, welche Mindestmenge an Informationen vorhanden sein sollte, um ein Gerät zu bedienen (z.B. Gerät und Schaltbefehl).

Interessant ist zudem, dass aktuell die Mehrheit eine Sprachsteuerung mit dem Auswendiglernen von Befehlen gleich zu setzen scheint. So wurde sehr oft gefragt: „Was darf ich denn sagen?“, dies verdeutlicht außerdem eine gewisse Barriere bei der freien, sprachlichen Interaktion mit einem technischen System. Daher wäre eventuell die Nutzung eines

Codewortes bzw. Namens sinnvoll, um das System anzusprechen und so für den Benutzer greifbarer zu machen.

Ähnliche Projekte lassen sich vorrangig zu Teilen des AISA-Projektes finden. Der Schwerpunkt liegt meist bei der einheitlichen, herstellerübergreifenden Anbindung von Hardware und damit verbesserter Interoperabilität (siehe z.B. iZ connected¹²). Dieses ist von hoher Bedeutung, da die aktuelle Situation markthemmend ist [HM09]. Fortschritte hier sind für das AISA-Framework von Vorteil. Andere Projekte mit beschäftigen sich mit der Entwicklung von Agententechnologien (z.B. DAI-Labor¹³), was ebenfalls durch die Zunahme verteilter Systeme von Bedeutung ist. Das AISA-Projekt verfolgt ähnliche Ziele, grenzt sich aber durch den Schwerpunkt der multimodalen, intuitiven Mensch-Maschine-Interaktion von diesen Projekten ab.

Literatur

- [GPZ05] Tao Gu, Hung Keng Pung und Da Qing Zhang. A service-oriented middleware for building context-aware services. *Journal of Network and Computer Applications*, 2005.
- [HM09] Reinhard F. Hüttel und Joachim Milberg. *Intelligente Objekte: klein, vernetzt, sensitiv*, Jgg. 1. Springer, Berlin, 2009.
- [SLP04] Thomas Strang und Claudia Linnhoff-Popien. A context modeling survey. *First International Workshop on Advanced Context Modelling, Reasoning And Management at UbiComp 2004*, 2004.

¹²Innovationszentrum "Vernetztes Leben - Connected Living", www.izconnected.de

¹³Das DAI-Labor forscht im Bereich der Smart Services und Smart Systems und vereint sechs Kompetenzzentren, www.dai-labor.de