DISSERTATION

Adaptive Behavior Arbitration for Mobile Service Robots in Building Automation

Submitted on the Faculty of Electrical Engineering and Information Technology, Vienna University of Technology, in partial fulfillment of the requirements for a degree of Doctor of Technical Science

Under supervision of

O. Univ. Professor Dipl.-Ing. Dr. techn. Ing. Dietmar Dietrich
Institut für Computertechnik
Technische Universität Wien
A-1040 Wien, Austria

and

Associate Professor Ph.D. Takahiro Yakoh
Department of System Design Engineering, Faculty of Science and Technology,
Keio University
3-14-1 Hiyoshi, Kohokuku,
Yokohama 223-8522 Japan

by

Dipl.-Ing. Charlotte Rösener
9526062
Maargasse 34, A-1230 Vienna, Austria

Vienna, July 2007   _____

Kurzfassung

In Gebäudeautomation und Robotik zeichnet sich ein Trend zur vermehrten Integration von Sensoren ab, die, zu immer komplexeren Überwachungssystemen zusammengefasst, eine exaktere Beschreibung der zu überwachenden Prozesse mithilfe mathematischer Methoden erzielen können. Diese Entwicklung hat auch massive Auswirkungen auf die Entwicklung neuer Methoden in der Systemsteuerung. Die folgende Arbeit beschreibt ein Funktionsmodell, das versucht, das autonome Verhalten von technischen Steuerungssystemen mithilfe der Nachbildung interner Mechanismen des menschlichen Verstandes auf Basis von Kenntnissen der Psychoanalyse und der Kognitionswissenschaften zu verbessern. Im Gegensatz zu anderen Forschungsarbeiten, die versuchen beobachtbares, intelligentes Verhalten direkt ohne Kenntnisse mentaler Funktionalitäten nachzubilden, ist das Ziel dieser Arbeit die Entscheidungsfindung des menschlichen Verstandes in seiner Funktionsweise zu beschreiben, um damit Entscheidungsfindung und autonomes Verhalten zu erzielen.

Auf Basis einer umfassenden Diskussion über die Einflüsse von Wahrnehmung, semantischem Vorwissen, inneren Konflikten, Gedächtnisleistungen auf den Entscheidungsprozess wurden Funktionen zur Entscheidungsfindung entwickelt und in einem Modell beschrieben. Um qualitative Aussagen über die Effektivität und Richtigkeit des Modells machen zu können, wurde eine spielähnliche Testumgebung geschaffen, in der zwei Gruppen Roboter mit gleichen physischen Eigenschaften aber unterschiedlichen mentalen Fähigkeiten um die limitierten Energiereserven konkurrieren sollen. Die Ergebnisse der Simulation zeigen dabei eine deutlich höhere Überlebenschance für Roboter mit höheren mentalen Fähigkeiten.

Abstract

In several application fields e.g. control networks and robotics, there is a trend towards installation and integration of massive amounts of sensors. With the increasing complexity of these systems new concepts in decision making for automation control become necessary. This thesis describes special issues of a functional model for automation control in order to improve autonomous behavior of a technical system. The model is based on functional theories of the human mind described by psychoanalysis and cognitive science. Unlike approaches based on simulation of intelligent behaviors, this thesis tries to focus on the certain internal functions of the human mind of which intelligent behavior may emerge.

Based on a comprehensive discussion about the roles of perception, emotional intelligence, memorizing experiences and competing instances for decision making, mechanisms are defined as functional blocks in an abstract control model. In order to allow qualitative conclusions about effectiveness and correctness of the model, a control architecture based on the proposed concepts has been configured for simulation. For evaluation two groups of uniformly embedded agents with different capabilities in autonomous control compete in game-like scenarios trying to take advantage of limited resources in order to sustain as long as possible. Results of this simulation give evidence that the emulation of higher mental functions like desire plans can extend the time of survival due to better adaptation to environmental conditions.

Acknowledgements

I would like to thank my advisor O. Univ. Professor Dipl.-Ing. Dr. Dietmar Dietrich for his sage guidance and support over the years. Without his unremitting efforts I would not have been able to complete this thesis. Thanks to his inspiration and advice that this research emerged from crude ideas at the beginning.

I would also like to thank my second advisor Associate Professor PhD Takahiro Yakoh for his great endeavors during my stay at Keio University, where I had the great chance to work with him and his students for two years. He has provided me with insight and direction since my start as a research student at his laboratory. Thanks to his endeavors I was able to broaden my horizons not only in the scientific area, but also in cultural matters.

My gratitude goes to Mark Solms, Edith Brainin and Patricia Pörscht for taking time for gripping and fruitful discussions. Their sage guidance has made psychoanalysis accessible to me, and left a deep impression in my scientific work and in my personal life.

My colleagues especially the members of ARS group at the Institute of Computer Technology deserve all my thanks and gratitude for being such a great team in this research project, always there to lend a helping hand and good humorous advice. With their discipline in communication and will for cooperation we made the world a little bit smaller doing research together although located in different parts of the world. Especially in the end of my thesis, they offered great encouragement and entertaining distraction (climbing somewhere deep in the mountains) helping me to find a healthier balance between intensive work and rest.

My parents receive my deepest gratitude for standing beside me with their rock-solid support which I can always rely on, granting me to live at ease. My father's way of life incorporates a lot of lessons that I start to understand. Special thanks to my uncle Roland for showing me what is important in life and how to fight for it.

Love and thanks to my boy-friend for all his patience, waiting quietly and bearing hard times in a long-distance relationship for years.

I would like to acknowledge all my close friends here and all over the world for their great understanding, when coming late from my laboratories, hiding behind work over weeks, or receiving answers with huge delays. Their words at late hours comforted me many times. It is great to know that they trust me and take me as I am. They are like a family to me.

"There is nothing more powerful than an idea whose time has come."

(Victor Hugo)

Vorwort

Die Arbeit beschäftigt sich mit einem neuen Funktionsmodell zur Systemsteuerung, wie sie beispielsweise für Gebäudeautomation oder Robotik benötigt wird, um deren Einsatzfähigkeit zu erweitern und deren Effizienz in der Problemlösung von unvorhergesehenen Situationen in verschiedenen, nicht vorgegebenen und veränderlichen Umgebungen zu verbessern. Das Modell kann als Vorbild für verschiedenste Steuerungssysteme dienen und darüber hinaus als Softwarearchitektur, z. B. in zukünftigen Generationen von mobilen Servicerobotern, implementiert werden, die voraussichtlich in moderner Gebäudeautomation benötigt werden. Mithilfe neuartiger Konzepte zur Prozesssteuerung, die in dem Funktionsmodell beschrieben werden, soll es möglich sein, einem autonomen System, z. B. einem Roboter, Aufgaben zu erteilen, die es autonom ohne vorgegebenen Lösungsweg erfüllen kann. Hierzu muss die Prozess- bzw. Systemsteuerung fähig sein, die gegebene, unter Umständen neue Situation zu evaluieren und aus ihr zusammen mit gelernten oder vorkonfigurierten Zusatzinformationen[1], die der Steuerung zur Verfügung steht, einen Lösungsansatz zu definieren, auszuführen und auf weitere Veränderungen während der Ausführung zu reagieren bzw. die Lösungsstrategie darauf abzustimmen. Für die Evaluierung soll das Funktionsmodell als Softwarearchitektur zur Steuerung von mobilen Robotern in einer spielähnlichen Simulation verwendet werden.

Bei der Entwicklung des Modells wurde darauf Wert gelegt, nicht ausschließlich auf klassische Methoden der Steuerungen, z. B. Regelkreise, zurückzugreifen, sondern diese vielmehr durch neuartige Konzepte zu ergänzen und gegebenenfalls zu ersetzen, so wie sie bei menschlichem Verstand zu finden sind und durch die Psychoanalyse in ihrer Funktionalität beschrieben werden. Jüngste Theorien und Erkenntnisse der Psychoanalyse bzw. Kognitionswissenschaften bilden die Grundlage für die Entwicklung von Mechanismen, die der Verbesserung von Entscheidungsfindung dienen sollen. Eine grundlegende Voraussetzung dabei ist die Behandlung von großen Mengen an Sensordaten, die bereits mithilfe von Symbolisierung [Pratl 2006, S. 25] aufbereitet, d. h. zusammengefasst und interpretiert wurden. Durch diese Vorverarbeitung wird es möglich, die Datenmenge auf Symbole mit gleichwertigem Informationsgehalt zu verringern [Pratl 2006, S. 26]. Aufbauend auf dieser in vorangegangenen Projekten [Pratl 2006], [Bruckner 2007] realisierten Methodik geht diese Arbeit noch einen Schritt weiter: Neben erweiterten Evaluierungskonzepten dienen neue Funktionseinheiten des Modells zur autonomen Steuerung und Erfüllung von Missionen, der situationsbedingten Konfliktbewältigung und der selbstständigen Problemlösung. Dabei muss die auf diesem Modell basierende Softwarearchitektur einerseits schnell auf dringende Notsituationen reagieren, andererseits aber komplexe Zusammenhänge erkennen und Konsequenzen abwägen. Dieser Spagat zwischen rechentechnisch schnell verarbeitend und rechentechnisch aufwendig soll in dieser

---

[1] Diese Informationen sind teilweise direkt Teil der Funktionseinheiten des Modells oder auch Teil der anwendungsspezifischen Konfiguration. Sie beinhalten sowohl Information über kausale Zusammenhänge und „Erfahrungswerte" in der Arbeitsumgebung als auch Verhaltensmuster und Informationen über interne Zusammenhänge und Fähigkeiten des Systems selbst, beispielsweise grundlegende Eigenschaften und ausführbare Handlungen des elektromechanischen Körpers des verwendeten Roboters.

Arbeit mit dem Einsatz von drei verschiedenen Rückkopplungsschleifen bewältigt werden. Wichtig ist dabei die Fähigkeit eines Systems, aufgrund von abstrahierten Beschreibungen von Situationen diese zu kategorisieren und sein Verhalten darauf abzustimmen und gegebenenfalls zu ändern. Dies ist bereits der erste wichtige Schritt, der auch für spätere Lernfunktionen benötigt wird.

Diese Arbeit beschreibt wichtige Mechanismen des theoretischen Modells, das zum Teil im Zuge eines Grundlagenforschungsprojekts (Artificial Awareness System) entwickelt wurde. Ziel dieser Arbeit ist es, ein Universalkonzept der Entscheidungsfindung auf Basis von Wahrnehmung und Gedächtnis zu erstellen, das ausgehend von Theorien der Psychoanalyse und der kognitiven Wissenschaften spezielle (interne) Funktionen des menschlichen Verstandes nachzubilden vermag. Auf diesem Funktionsmodell basierende Steuerungen sollen neue Fähigkeiten zur selbständigen Aufgabenbewältigung hervorbringen, wie etwa Gliederung und Verteilung von Einzelaufgaben, Entscheidungsfindung und Erstellen von potentiellen Lösungsstrategien. Diese Fähigkeiten sollen in dem vorgestellten Modell durch folgende Funktionen ermöglicht werden:

- Die emotions- und triebgesteuerte Filterung und Evaluierung von wahrgenommenen Bildern und Episoden erlauben, Gegebenheiten des direkten Umfelds eines Systems mit dessen internem Zustand und Bedarf in Relation zu bringen und entsprechend zu berücksichtigen. Diese Methode ist eine wichtige Voraussetzung, um ein Gleichgewicht zwischen extern aufoktroyierten Aufgaben und den Bedürfnissen des Systems zur Selbsterhaltung zu schaffen.

- Drei verschiedenen Kontrollschleifen von unterschiedlicher Komplexität verbinden Leistungsfähigkeit und Präzision. Sie vereinen relativ konventionelle Konzepte einfacher regelbasierter Steuerungen mit erweiterten, rechentechnisch aufwendigeren Konzepten zur Planung und zu strategischen Problemlösungen. Interne Konfliktbewältigung- und Verdrängungsmechanismen durch teilweise komplementäre Verhaltenstendenzen schaffen zusätzliche Adaptionsmöglichkeiten.

- Verschiedene Gedächtnissysteme erlauben die Beachtung teilweise anwendungsspezifischer (sozialer) Regeln und die Adaption und Verbesserung der Steuerung durch Einbeziehen von individuellen Erfahrungen, die als Grundvoraussetzung jeglicher Form des Lernens angesehen werden.

Das Modell kann als Entwurf für eine Reihe von verschiedenen Steuerungssystemen betrachtet werden. Allerdings liegt der Fokus in der Steuerung von mobilen Robotern, da Mobilität und Verkörperung[2] als wichtige Voraussetzungen für die Anwendbarkeit mentaler Funktionen, wie sie in der Psychoanalyse beschrieben werden, betrachtet werden. Nachdem erwartet wird, dass sich Serviceroboter zukünftig als Verbindungsglied zwischen ausgereifter Situationserfassung und -erkennung und aktiver Hilfestellung im privaten Haushaltsbereich als überaus nützlich erweisen sollen, wurde eine eigens entwickelte Simulationsumgebung geschaffen, in der u. a. auch die zukünftigen Einsatzmöglichkeiten solcher Roboter in der Gebäudeautomation aufgezeigt werden.

---

[2] Diese Voraussetzung, dass ein System einen „Körper" besitzt, bedeutet nicht notwendigerweise, dass dieser exakt der Beschaffenheit eines biologischen Organismus zu folgen hat. Wichtig in diesem Zusammenhang ist die räumliche Ausdehnung und physische Begrenztheit des Systems, durch die eine klare Unterscheidung zwischen systemeigenen Elementen und fremden Elementen der Umgebung möglich wird.

Das in dieser Arbeit vorgestellte Funktionsmodell ermöglicht es Systemen, die man als autonome Agenten betrachten kann, sowohl effizient Aufgaben allein oder auch in Zusammenarbeit mit anderen gleichartigen oder heterogenen autonomen Agenten (autonomen Systemen) zu übernehmen als auch mit einem übergeordneten Steuersystem, wie oft in der Gebäudeautomation zuoperieren die damit verbundenen Teilaufgaben erfolgreich und autonom auszuführen.

Anhand von Simulationen soll die Tauglichkeit dieses Modells diskutiert und untermauert werden. Hierfür wurde eine repräsentative spielähnliche Umgebung geschaffen, in der Roboter (verkörperte Agenten) bzw. Gruppen von Robotern mit verschiedenen komplexen Situationen konfrontiert werden, in denen eine möglichst profitable Entscheidung (für ihn bzw. für sein Team) zu fällen ist und in denen entsprechenden Maßnahmen effizient durchzuführen sind. Auf diesem „Spielfeld" treten Roboter mit gleichen physischen Eigenschaften, aber unterschiedlich ausgefeilten Steuerungssystemen jeweils als ein homogenes Team gegeneinander an und müssen in der Gruppe auf Probleme reagieren, Konflikte lösen und situationsbedingt entstehende Aufgaben übernehmen. Die Gruppe, die mit Problemen am effizientesten umgehen und diese lösen kann, wird in der Simulation erfolgreicher agieren können und dem Spielziel, nämlich dem Erhalt der Funktionsfähigkeit der Gruppe, am nächsten kommen. Anhand dieses Wettbewerbs der zwei Gruppen wird der erfolgreiche Einsatz von auf dem Modell basierenden Steuerungen bewertet.

Der Aufbau dieser Arbeit lässt sich wie folgt beschreiben (die Abhängigkeiten der einzelnen Kapitel sind in Figure 1 dargestellt): Das erste Kapitel gibt eine allgemeine Einführung in die Thematik wieder und beschreibt die Erwartungshaltung bezüglich zukünftiger Entwicklungen und Probleme. Diese wiederum führen direkt in das zweite Kapitel über, das sich vor allem mit den Anforderungen und Randbedingungen beschäftigt, die sich im Zuge der zu erwartenden Anwendungen ergeben. Im dritten Kapitel wird eine umfassende Begriffssammlung in Verbindung mit verschiedenen Konzepten aus allen drei Disziplinen (Gebäudeautomation, Robotik und Psychoanalyse) wiedergegeben, die die wichtigsten Aspekte und Mechanismen anführen und beschreiben, die für die Ausführung dieser Arbeit ausschlaggebend sind. Ihre wichtigsten Eckpunkte sind maßgeblich für das Design des Modells, das im sechsten Kapitel beschrieben ist. Aufgrund von Kapitel 3 wurde eine generelle Vorgehensweise und Methodik in Kapitel 4 definiert, nach deren Annahmen in dieser Forschungsarbeit vorgegangen wurde. Im fünften Kapitel werden nochmals die wichtigsten und teilweise vorangegangenen Projekte, die den stärksten Einfluss auf diese Arbeit genommen haben, vorgestellt und diskutiert. Kapitel 3, 4 und 5 resultieren im Design eines abstrakten Funktionsmodells für Steuerungssysteme, das in seinem Aufbau, seiner Funktionsweise und seinem Ablauf im Überblick (wie im Detail) in Kapitel 6 vorgestellt wird. Um die Wirkungsweise und Tauglichkeit des relativ abstrakten Modells für die Steuerung von beispielsweise mobilen Robotern anhand konkreter Anschauungsbeispiele wiederzugeben, wurde eine eigene Simulationsumgebung geschaffen, die in Kapitel 7 beschrieben wurde. Im letzten Kapitel dieser Arbeit werden nochmals die wichtigsten Ergebnisse zusammengefasst und ein Ausblick auf folgende Forschungsentwicklung wird angerissen.

Preface

The approach deals with a novel functional model for system control used e.g. in building automation or robotics to widen their application field and improve their efficiency in task allocation in unstructured, changing environments. The functional model can be used as an archetype for different control architectures e.g. in robotic control as it is expected to be used in building automation. Due to new concepts in control described in the model, autonomous systems, e.g. robots, shall be capable to complete assigned missions autonomously by recognizing and evaluating new situations via environmental and internal information[3], in order to create an action plan and adapt the plan during the mission.

For the design of the model theories of psychoanalysis and other cognitive sciences describing functionalities of the human mind have been used besides classic concepts of automation control. Resent findings that shall explain the inner functions of the human mind are used for the design of more capable methods for behavior arbitration. Based on data processing functions of succeeding projects ([Russ 2003], [Tamarit 2003], [Pratl 2006], and [Bruckner 2007]) that allow the filtering and compression of massive sensory data via symbolization, this approach provides additional functions for mission completion, situation-dependent conflict resolution and autonomous problem solving. The software architecture has to have high-performance reaction in emergencies and be capable of recognizing complex constraints and considering potential consequences caused by different actions. This difficult combination of high performance and complex control shall be covered by three different control systems. Therefore, the robot's capability to categorize situations using abstract templates of situations (so-called abstract images) is crucial and meets the requirements for learning mechanisms.

The proposed work presents special mechanisms of a general theoretical model that has been developed as a part of a basic research project (Artificial Recognition System). The universal control concept is derived from findings in psychoanalysis and cognitive science. The goal of this approach is to integrate novel mechanisms emulating unique capabilities of the human mind to overcome problems in current solutions. Control software based on the function modules of this model shall be able to complete missions autonomously covering important capabilities like task composition, decision making, and action planning. The model can provide these functionalities based on the following concepts:

- Emotion based filtering and evaluation of images and episodes provide goal directed behavior balanced between applied missions and the needs for self-preservation of the system. They bring internal (bodily) needs in context to the perceived environmental images and stored, "experienced" episodes.

- Three different control modules of different hierarchy bridge the gap between high performance and accuracy. They unify conventional, rule-based control enhanced with concepts of mission planning. Conflict resolutions and inhibition allow flexible behavior adaptation.

---

[3] The information originates either directly part from functional entities of the model, or is related to the application dependent configuration. The information inherits causal connections and "experimental values" about the environment and constrains based on the abilities of the system itself, e.g. basic properties and executable actions of the robotic body.

- Memory systems inherit application dependent world knowledge and allow the storing of individual experience, a crucial pre-condition for learning in general.

The model can be used as a blue print for a variety of control systems. However, the main focus of this model is to improve autonomous control systems of mobile service robots, as embodiment is seen as a major requirement for the validity and applicability of theories provided by psychoanalysis. Based on the assumptions of the development in building automation, where service robots play the missing link between observation and active support, a simulation environment has been built, where robots armed with model based control systems are situated in.
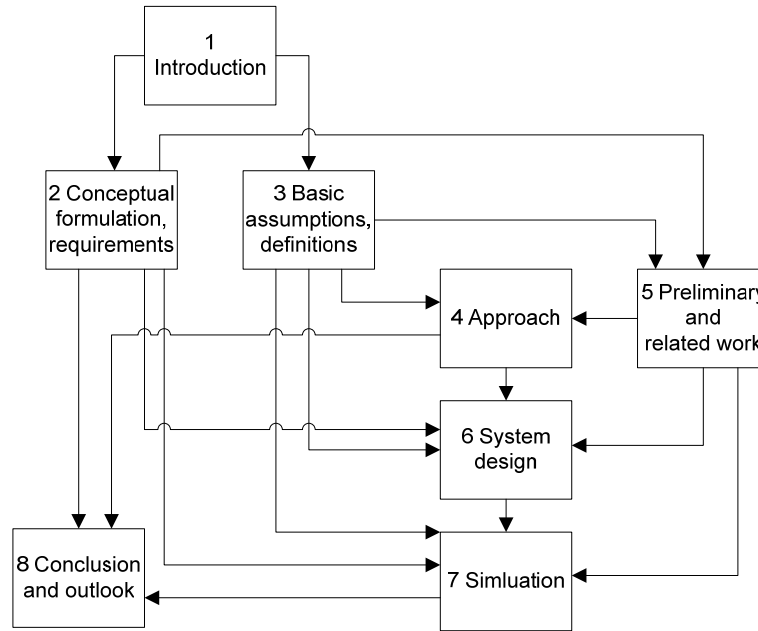

Figure 1: Structure of the thesis

The proposed model also allows the efficient task decomposition in cooperation with heterogeneous robot teams or control systems. In simulation, the potential of model-based robotic control is tested and discussed. For that reason, a game-like environment has been designed in which robots compete in different groups and try to find an optimum strategy in diverse (unknown) situations: a group's success is determined by deliberating and choosing appropriate behavior within the group, including role assignment and task allocation in a given setting. The result of this simulation provides a measurement tool for qualitative evaluation.

The thesis is structured as follows (Figure 1): the introduction gives a general overview of the background and objective of this work. The latter will be specified in more detail in Chapter 2, which gives a detailed overview of requirements that are in the context of this approach. The propositions put forward in Chapter 3 outline a comprehensive set of important mechanisms and introduce all necessary guidelines that are essential to design the model, which will be looked at in greater detail in the design rules of Chapter 4. Presenting ongoing preliminary work and findings of other disciplines in Chapter 5, this will detail the major influences and preceding solutions of this work. In Chapter 6, a general over-view and special aspects of this model will be explained. This very abstract model shall be explained further in a concrete simulation example giving evidence of its applicability. The conclusion will sum up the results and give an outlook for further development.

Table of Contents

1 Introduction

Demographic transitions can be found in most highly developed industrial nations, e.g. in the majority of European countries, but also in Japan and in some states of the U.S., where the age pyramid has gone through dramatic changes. Improved medical care extends the average lifespan, entailing a significant increase in elderly people within the society, which faces different needs especially in context of successful aging. This well-funded group of buyers will become an important factor in the future economy, which will be spurred to provide age-based products to meet the needs of this increasing part of the population. Especially health and long-term care require further improvement in technologies such as service robots. Manpower in nursing and rehabilitation cannot meet theses needs alone and is overloaded due to a decline in new employees because of a dwindling next generation. Acceptance of automatic (robotic) assistance to restore the physical performance of disabled persons either in long-term situations or during recovery is crucial [Taylor 2006]. This also requires modern designs of domestic buildings to incorporate all facilities that ease the daily life and offer elderly and handicapped people an independent existence without external (manual) help. To achieve these objectives, building automation technologies combined with solutions in other technologies, like robotics, are inevitable [Pratl 2007].

The majority of technologies in building automation are founded on control networks, programmable networks of computers and electronic devices that can among other functions monitor and control the mechanical and lightening systems in a building. However, as [Kabitzsch 2002, p. 17] proposes, it is more appropriate to use a more abstract function-based view of these systems. How will the modern home or office look like in the future? Due to the durability of buildings utilized for many years, this question has to be resolved without delay. In the last decades, building automation has gained importance by becoming an established factor supplementing modern life [Pratl 2006, p. 1]. The goal is to create a flexible and adaptable building and reduce energy and maintenance costs, but also to improve comfort and feasibility. Providing multi-functional systems shall give buildings a new range of applications, e.g. security, monitoring and access authorization, safety and comfort (including lightening and HVAC[4] systems), energy management, control of white goods[5] and brown goods[6] etc. [Kastner 2004], [Kabitzsch 2002, p. 21-27]. In future, these applications will be enhanced with a main focus on new functionalities like cleaning, nursing and rehabilitation by exploiting synergy effects of existing infrastructure.

With new technologies, e.g. ambient computing or smart personal objects technology, building automation systems have broadened their application dramatically [Pratl 2007]. In contrast to former objecting to facilitate production halls and other functional buildings with a higher level of automation, the application field focuses nowadays appreciably on office buildings and domestic use to improve modern living conditions. Within this research field a paradigm shift occurred. Fieldbus systems, as they are used for building automation, can be only the first step [Loy 2001, p. 2]. Due to the increasing

---

[4] Heating, Ventilation and Air Condition
[5] This term includes all types of household appliance, e.g. dish washer, washing machine, etc.
[6] This term includes all forms of home and entertainment electronics, e.g. television set, stereo set, etc.

numbers and actuators a higher and more efficient level of data processing is required to be able to cope with the massive data reaching an optimum of performance [Pratl 2007], [Dietrich 2007], [Kastner 2004], and [Kabitzsch 2002]. In the case that a building automation system does not only collect and analyze information about its state, but also reacts based on the results of the analysis, there must be new mechanisms to do so in a wider range [Pratl 2006, p. 1]. This is the application field of service robots in the future. Service robots represent the missing link between intelligent data analysis and active support for people in their everyday lives [IFR 2006, p. 417]. Especially in an over-aged society, where an increasing number of people become physically incapacitated and cannot operate independently due to age or ailment, the demand for increasingly sophisticated control networks in modern living spaces becomes apparent. Due to several different tasks, the shape and ability of robots will differ, but the basic challenges of autonomous task allocation, fast situation recognition and behavioral decision making without external input stay the same.

1.1 Service robots today

The World Robotics 2006 [IFR 2006] contains amongst others the 7th comprehensive market survey of service robots carried out by the IFR[7] Statistical Department by order of The United Nations Commission of Europe (UNECE) in [UNECE 2004], [UNECE 2005], and [IFR 2006]. Due to a surge of investments in industrial robots in the last three years, and further boosted by plummeting robot (relative) prices in 2005, which were on average about 23% of those in 1990 without quality adjust [IFR 2006, p. 125], the industry is booming. Although the prices of robots are falling rapidly, the purchase and handling of robots still needs improvement. Especially industrial robots and service robots for professional use cannot be acquired simply "over the counter". Besides their industrial counterparts, service robots gain more and more attraction. Although service robots have a small market share in the last years [UNECE 2005] when compared to industrial robotics, the predicted potential of non-industrial applications exceeds those of industrial ones by far [Prassler 2000], [Schofield 1999]. However, the often proclaimed great rush on robot-aided domestic applications has been slow to start [UNECE 2005]. With the end of 2005, up to 31,600 units for professional use have been installed (which is 3.4% of 923,000 installed industrial robots by the end of the same year) [IFR 2006, p. 382]. Nevertheless, great advances have been made in the field of robotics during the last decade, which have heralded reversal trends in the domestic use: according to [IFR 2006] a world stock of 1,900,000 units will exist by the end of 2005. More flexible types of multi-purpose robots are projected to emerge in the next years, forming a trend from expensive industrial devices to small-sized, light-weight and low-cost systems ready for the private market [Prassler 2000]. Other than their industrial counterparts, which normally fulfill predefined tasks by operating in structured and adapted working cells separated from human operators, service robots face a higher level of insecurity in environment and applied tasks [Bolmsjö 1995]. Furthermore, service robots are supposed to work autonomously and over long periods of time without human control [Prassler 2000]. To overcome these technical, social and economical challenges will be the key factors in the predicted great rush on service robots in the next years.

---

[7] International Federation of Robotics

Still, the majority of robotic research concentrates on professional service that can be categorized into the following fields of application: field robotics, professional cleaning, inspections systems, construction and demolition, logistic systems, medical robotics, defense, rescue and security applications, and underwater systems. Especially underwater systems (18% or 5,680 units), followed by cleaning and laboratory robots (17%), and defense, rescue and security applications (16%) represent the main categories with the highest numbers of installed units by the end of 2005 [IFR 2006, p. 378]. The prices of these types are varying significantly with their purpose, and especially defense, rescue and security applications and medical use require large investments per unit. Therefore, the highest value accounted for all installations can be found in defense, rescue and security applications due to the high costs per robot followed by underwater systems, where a high number of installations can be found [IFR 2006, p. 380].

1.9 million robots for domestic use and about 900,000 for the entertainment industry were installed and in use by the end of 2005. In [IFR 2006, p. 382] IFR projects 3.9 million units of new installed service robots for domestic use in the next three years. The unit value of a robot in domestic and personal use has to be a fraction of these types of robots. In contrast to robots for professional use, these robots have to be produced for the mass market at reasonable prices to become accepted products. At present, a comparatively limited area of application fields are vacuum cleaning, lawn-moving, and home entertainment including hobby, education and training systems. The stock of handicap assistance is still small but IFR [IFR 2006, p. 380] states the expectation that this will double in the next four years as an alternative for the care of the elderly and handicapped by individuals. The acceptance of robots in these applications is low as robots are still claimed to be too clumsy to perform in an ordinary household as a major support in everyday life. This might be one of the main reasons why the majority of robot manufacturing still concentrates on industry robots or service robots for professional use where robot-aided applications are required. In households, service robots will not only clean floors or mow lawns but they will also assist handicapped people and will be used for other applications, e.g. personal transportation, home security or surveillance [IFR 2006, p. 380].

1.2 Expected potential of service robots in building automation

Definition 1.1: "Rehabilitation is an activity which aims to enable a disabled person to reach an optimum mental, physical and/or functional level." [Bolmsjö 1995]

Robots for domestic use are still regarded as niche products for professional services [IFR 2006, p. 386]. Compared to professional service, robots have to obtain a similar efficiency in low-price solutions for domestic services [IFR 2006, p. 416]. Building automation and robotics for rehabilitation shall provide physically disabled persons with the necessary tools to support themselves in their daily activities and improve their living and working conditions [Bolmsjö 1995]. Robots shall liberate people from tiring, monotonic work by operating by themselves [IFR 2006, p. 416], but also assist in situations where the presence of individuals is too dangerous (hazardous environments) or physically impossible (undersea, narrow places, high altitude). This requires the development of more flexible systems in unstructured environments with a focus on the necessary functionalities as defined and evaluated by users [Bolmsjö 1995].

Current control networks in building automation are highly distributed systems, containing a vast number of sensors and actuators of varying complexity [Kastner 2004]. Conventional architectures follow a centralized design, integrating the computational power centrally and exploiting the properties of simple wiring to a lesser extent [Loy 2001, p. 3]. However, extensive data transfer and transformation has become problematic with an increase in complexity and amount of data [Loy 2001, p. 3]. To cope with the high amount of information, new decentralized approaches combined with new methods of data processing and data management with increasing computational power have been introduced. The vast number of information nodes found in modern buildings seems to be only the beginning of a far higher level of information processing in the future [Pratl 2006, p. 13]. Besides elementary sensing, which is a premise for any cognitive process, the evaluation of situations in different environments, e.g. at home, in an office building etc., is necessary. This was one of the reasons why a new concept of symbolizing sensory data was introduced in the preliminary works of [Russ 2003], [Tamarit 2003], [Pratl 2006] and [Bruckner 2007] and under the direction of o. Univ. Prof. PhD Dietmar Dietrich. Nevertheless, the provided actuators of these systems appear comparably static, giving the control systems a very limited range of possible actions. Further concepts and tools are inevitable to select and execute proper actions. A new generation of mobile service robots has to be integrated, completing systems of building automation to reach a new level of automation.

In building automation, service robots and individuals are integrated in a common work space where people are part of the process and task execution [Bolmsjö 1995]. Service robots in domestic use live with humans in human societies. As these robots are in direct contact with people in their living space, additional aspects, e.g. safety, are inevitable and represent the missing link to take actively part in the procedures occurring in a house. Service robots can be the (autonomous) actuators of a control network. In these unstructured and changing environments robots have to understand the environment they are situated in rather than simply collect data, as they have to take action autonomously in a wider scope than before. This is of particular importance since in general there is no person specially trained in the technology but people of low or no interest and knowledge in programming with physical problems that need maintain the devices in case of an error [Bolmsjö 1995]. To ensure the survival of the service robot, to remain intact and able to fulfill tasks, and for the safety of others (human or robot), these types of robots need higher level strategies and reasoning than are provided so far. In order to interact with changing environments that are provided in human life, technical systems have to undergo a similar evolution as biology has done before. The human archetype is a successful example of how to act in unknown environments. Especially the ability to not only react on given inputs but also plan future behavior upon predictions, based on the experience of past situations, is a powerful and highly desirable capability. This requires unique functionalities underlying all higher mental abilities, which have been unachievable in the technology so far.

Provided with a combination of newly designed methodologies, which can be adjusted by learning processes, robots shall become capable of behaving as autonomous, embedded agents selecting and reaching goals autonomously instead of a blind following of machine instructions given by any kind of supervision. Therefore, the robot must be capable of choosing and planning its own strategy, determined by the needs of its structure as well as by preconditions to fulfill the tasks of a mission and the needs of interacting with its environment.

The goal of this research is to facilitate robots and control system with new methods to cope with the problems facing traditional Artificial Intelligence (AI) based on sensing-modeling-planning-acting. According to [Mochida 1995], [Singh 2003], and [Ratanaswasd 2005] the main concerns of AI are:

- Flaws in continuously changing environments: for applications with a wide range of actions under ever changing circumstances, the main challenge in control architecture is not to record the exact state of a robot and its environment, but possess basic knowledge about causalities. This allows also appropriate predictions for arranging action sequences to increase performance.

- Focus on attention: the unfiltered amount of data collected by sensory systems during a journey through the working space requires high computational effort, which leads necessarily to longer calculation times.

- Automatic resource allocation: there must be additional probability estimation for emphasizing the environmental changes and adapting mission plans.

These problems have led to more centralized control designs that require external computation of decisions due to limitations of mobile robotics so far [Mochida 1995]. Although these concepts show elegant examples of robot mapping and task learning suitable for many applications, centralized control systems where robots act according to external command entail problems in complexity, requiring extensive computing, and risking late responses in dynamically changing situations [Jones 2005].

1.3   Research background

One of the main research areas of the Institute of Computer Technology in the context of building automation is the design of future control system using existing and new technologies. O. Univ. Prof. Ph.D. Dietmar Dietrich has already predicted in [Dietrich 2000] that although classic engineering and computer science have contributed in evolution so far, problems due to increasing complexity demanding flexibility and scalability remain unsolved. To achieve these goals, new concepts in building automation founded on biological systems are required [Dietrich 2000]. In the past years, a number of different research projects have been introduced that touch on this topic. Especially with the foundation of the "Artificial Recognition System" (ARS) project in 2003, which concentrates on information processing concepts emulating unique capabilities of the human mind, and the forming of a research group to build a team for robot soccer, creating small-size mobile robots, new opportunities of interdisciplinary work in designing systems with a wider range of application fields have arisen.

In the first approach [Russ 2003], which deals with situation recognition, case-based reasoning is used as a starting point. Based on a similar idea as shown in the model of [Aamodt 1994], a more flexible handling of unforeseen events has been achieved in order to avoid the repeated processing loops of small details. This model has been deployed to achieve a higher level of automation and "intelligent" behavior within building automation in the "SmartKitchen" project. Another goal of this project was to find a solution to cope with massive sensory data in order to facilitate fast decision-making processes. One of the major problems so far is that there is no hardware infrastructure that can provide and prepare the massive sensory data to be expected in the future. This limitation leads to the extensive use of simulations for evaluation. This pioneer project has carried over to the "Artificial Recognition

System" (ARS)[8] project, which focuses on two major aspects relevant for future systems in building automation: one goal is to build up an advanced perception system based upon bionic models, which shall be capable of recognizing complex processes within a building by means of sensor fusion of massive sensory data and symbolization. This research is conducted in a separate project called ARS-PC (Consciousness), where [Pratl 2006] succeeds with the introduced concepts of [Russ 2003] and [Tamarit 2003]. The goal of ARS-PC [Pratl 2005a] is to find control concepts that are capable of recognizing situations considering a massive amount of redundant sensory data and extracting essential information [Pratl 2006, p. 2].
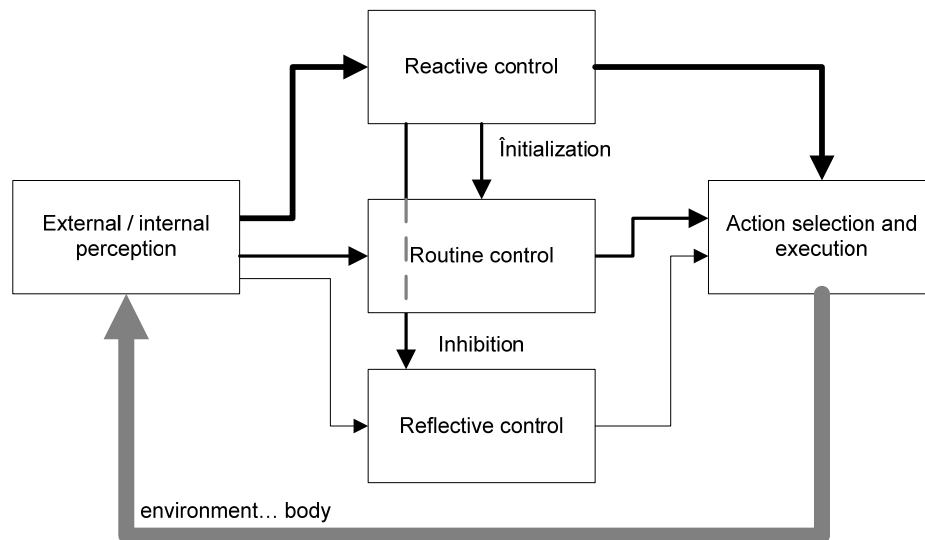
Figure 1.1: Three control loops of behavior architecture

So far, the ARS system has been a supervisory system with very limited potential for taking action itself and depends on human operators for acting. The second objective is to enhance the control system by introducing cooperation with a new generation of mobile service robots to broaden the action range of automatic control. This is part of the ARS-PA (PsychoAnalysis) project, which concentrates on control architectures for mobile robots with focus on the special requirements for service applications in domestic area. The behavior model-based software architecture is based on new concepts inspired by psychoanalysis and other disciplines in cognitive science. The control model shall allow independent action selection and task allocation, which are crucial capabilities for diverse service applications. It shall facilitate the robot to work autonomously, in team with other service robots and in cooperation with the mainframe of control systems.

Both parts of this project are developed to operate independently in order to allow for the adaptation of existing systems. In the meantime, the ARS project has split up into further future subprojects, e.g. the BASE (Building Assistance system for Safety and Energy efficiency) project, which provides a self-learning system that can learn what is regarded as normality and will alert in the case of deviations [Bruckner 2007]. This project is coordinated by ARCS Seibersdorf Research GmbH - Geschäftsbereich Informationstechnologien.

Besides building automation, the second research area deeply involved with this concept is robotics, which has been introduced so far. Within the Center of Excellence for Autonomous Systems at the

---

[8] Detail can be found in [Pratl 2005b].

Vienna University of Technology, members of the Institute of Computer Technology have developed generations of a tiny, fully autonomous mobile robot, called Tinyphoon [Novak 2005], a successor of Roby-Go [Novak 2004a], both used primarily for robot soccer. With all its components, the robot Tinyphoon fits into a cube with an edge length of 75 mm. Although it has been built for playing robot soccer, efforts have been made from the start to keep the design as open as possible so that the robot can easily be used for other applications [Mahlknecht 2004a], [Novak 2004b], and [Egly 2004]. The main innovative goal of the design of the tiny robot is to enable it to act completely autonomously within its environment. This mobile robot has been designated for the first implementation of a derivative of this control architecture in the near future.

Within this thesis, an abstract architecture for behavior control of all types of autonomous agents, with focus on service robots, is used within multi-robot systems and other control systems. Therefore, a general behavior model has been designed as a composition of functional modules. The behavior-based architecture shall facilitate a mobile robot to perform missions independently, containing a series of sub-tasks by selecting the appropriate action. The main focus of the model is the selection of strategies to avoid and solve conflicts by adaptive action selection due to current environmental conditions, including the activities of other robots and the internal state of the robot itself.

The main work of this thesis is embedded in the research projects of the Institute of Computer Technology and founded on previous approaches. Furthermore, this research is the result of a research exchange from April 2005 to March 2007 on the Keio University, Yokohama/ Japan, where I have participated in projects of Associate Prof. Ph.D. Takahiro Yakoh and his laboratory to gain further understanding in particular of aspects in robotics, information theory, distributed systems, and real-time communication.

The designed and proposed model of this research is kept abstract to provide a wide range of capabilities for multi-purpose architectures in future systems. The controll architecture shall give new concepts that are necessary for all service robots with similar difficulties. However, the idea is to keep the focus on specific applications emphasizing precise requirements.

2 Conceptual formulation and requirements

Robots shall help to reduce the need for manpower where skilled labor is in short supply or where the application is too dangerous for people. However, traditional designs of work spaces and organizational procedures have been mainly technology–driven, focusing on the company's employees as data-processing systems which can be replaced rather than supported in their tasks [Lueg 1997]. Considering the flaws of these concepts, the demand for a more user-specific view on service-providing systems has emerged [Bolmsjö 1995], and [Lueg 1997]. Although originally classic applications have been used for undersea or air and space tasks, service robots shall be installed for household maintenance or other applications in buildings in order to assist in the improvement of living conditions by taking over tiring tasks.

2.1 Application dependent requirements on service robots

All service-oriented applications have the common requirement of appropriate behavior in a highly dynamic and unpredictable environment. A robot used in these applications has to react reliably, robustly, coherently and flexibly to these changes. This is one of the unique capabilities of the human mind: to find new solutions for complex problems of several consisting constraints, which allow for the planning of future actions and shaping of the environment [Ratanaswasd 2005]. But what is the advantage in giving robots human-like behavior? And do we need the complexity of a human mind for these applications? With a broader scope of functionalities and duties, robots have left their usual domains in factory sites and start to "explore" humans' living space, interacting with humans, and assisting them in their tasks. This changes also the relationship between man and machine. Systems like these shall become partners in work rather than tools for humans. Therefore, it is a necessity to make robots *understand* their environment in order to act in it without extra supervision. Even current advanced systems still need the support of a human advisor who complements the system in what it lacks: a common sense to evaluate and predict real situations, mostly in interaction with other people. But in order to make robotic systems acceptable in the living space, they have to offer a high benefit in complementing the abilities of humans, no matter if they are physically sound or have disabilities, and in cooperating in tasks that might be impossible without robotic assistance [Taylor 2006]. But at the same time, cooperation means responsibility for both the human and the robotic side, and requires regulations. In general, there are three rules a system has to obey when cooperating with human beings. It should provide capabilities close to the following guideline[9] [Schraft 2004, p. 30]:

(1) A system (tool, robot or agent) should provide safe use. In case of an autonomous system, the system must not harm a human being or allow that humans get harmed through inaction.
(2) A system (tool, robot or agent) has to perform correctly unless this means a threat to users. In case of an autonomous system, the system must obey commands given by human beings as long as these orders do not conflict with the first rule.

---

[9] from Isaak Asimov's "laws of robotics" listed in the short story "Runaround", in 1942 [Schraft 2004, p. 30]

(3) A system shall remain intact as long as possible unless its destruction is necessary for safety. In case of an autonomous system, the system has to protect its own existence as long as the protection does not conflict with the first and the second rule.

Although these rules originate from fiction, they show a very user-oriented view of the system requirements, which is actually desirable for service applications as these systems do interact with technically unversed users [Schofield 1999].The model described in this thesis focuses on autonomous mobile robots especially for services. The basic idea is that robots shall assist in many household domains and thus gain higher acceptance. Besides crucial assistance in rehabilitation, robots can take on further tasks once other control systems have been completed. Typical and diverse operations that are carried out in domestic applications are [IFR 2006, p. 419-420]:

-   Cleaning and other domestic tasks: robotic assistance frees individuals from tiring tasks, taking over routine jobs, e.g. vacuum cleaning, lawn mowing, window cleaning, etc. They might be slow but can execute a given task all day long [IFR 2006, p. 416] as long as they do it with the necessary accuracy. Special requirements are autonomous path planning and execution [Schofield 1999], autonomous recharge and energy control [Schraft 2004, p. 15-23].

-   Health care: personal robots, including robotized wheelchairs, can assist handicapped people in their living and work space, monitor the condition of the assisted individual, and call for help in the case of an emergency. This requires capabilities in payload [Bolmsjö 1995], safe navigation in peopled environments, circling and guiding people around unexpected obstacles, monitoring of individuals and their vital functions, etc. [IFR 2006, p. 420]. In particular health care requires multi-purpose robotic systems [IFR 2006, p. 421].

-   Home security and safety: to complement other systems, robots that act as safeguards as protection against intruders require capabilities for the identification of individuals and situations [IFR 2006, p. 419].

-   Education and entertainment: systems can use mobile robotic systems to provide advanced interfaces interacting with humans in providing mobile communication, information platforms, networking, etc. [IFR 2006, p. 419]

When using robots in these application fields, specific requirements have to be fulfilled under more challenging conditions, as resources of energy, computational power, etc., are more limited on a mobile robot [Novak 2005]. Therefore, these systems have to meet further objectives:

-   Robustness: service robots have to adapt to uncertain situations in changing environments [IFR 2006, p. 386]. Therefore, the navigation and mission complementation in a complex environment shall be warranted, even when there is little or no communication with other robots or no automation control system acquirable [Ueno 2000].

-   Reliability: robotic systems shall be a compensation for missing skilled labor and be of untiring assistance under a heavy payload when cooperating with humans whose safety might depend on the reliability of robotic assistance. Furthermore, robots are assumed to take over essential tasks in hazardous environments, and therefore cannot rely on human intervention in case of malfunctions [IFR 2006, p. 386]. These circumstances require a reliability which is better or at least equal to the one of physically able humans [Sugawara 1999].

- Flexibility: acting in continuously changing environments, where the robot has to interact with humans or other robots in order to complete a mission successfully, require the adaptability of behaviors when considering changes [Bolmsjö 1995], [Taylor 200], and [Schofield 1999].

- Sustainability: energy storage is very limited on a mobile robot. An efficient use of these resources with estimating strategies for optimization is highly desirable [Novak 2006].

- Safety: as robots may have to cooperate with humans during certain tasks, it is imperative to guarantee that these robots cannot harm human life [Schraft 2004, p. 28]. Human safety in particular is crucial in the case of a shared work space with humans. Figure 2.1 shows general aspects which have to be considered in this context in order to achieve safe designs in physics and control.

- Security: similar to the building automation system itself, a service robot has to fulfill basic security requirements, especially in communication with a mainframe or within the robotic team, to ensure the correctness and confidentiality of critical transmitted data, such as positions and identifiers of robots, commands and sensory data.

- Simple maintenance and handling: as it is expected that service robots work with untrained operators without extra technical support structures, their handling must be kept simple and their control should be sound and easy to remember [Schofield 1999], in order to give users more time on concentrating how the robot operates and executes assigned tasks [Bolmsjö 1995].
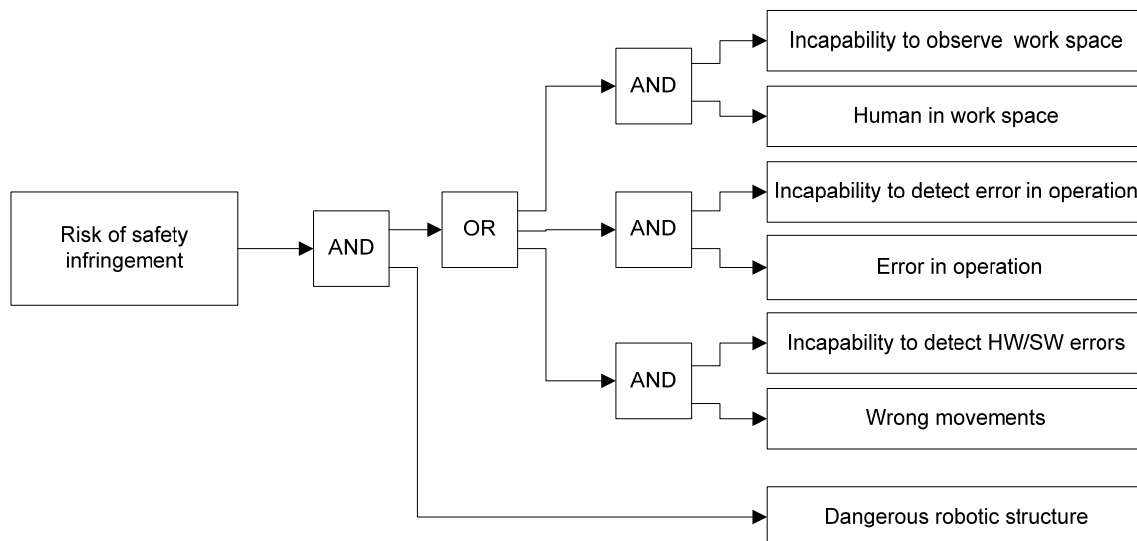
Figure 2.1: Design rules achieving safe robotic designs according to [Schraft 2004, p. 28]

Further objectives like low-cost manufacturing pose a significant challenge for the long term [IFR 2006, p. 386]. Further specializations, e.g. shape and physical properties, will be necessary. When solving more complex problems, domestic service robots must also be capable of cooperating with other robots in a team, but also assist humans in certain tasks. Well coordinated robot teams can operate more efficiently than single robots. Tele-operation of large robot groups appears to be communication-intensive and error-prone [Buffet 2001]. A service robot within a multi-robot system has to share global goals but has also to make sure it is capable of working efficiently. This can lead to conflicts that have to be resolved by the robot without external advice.

Due to the different capabilities for human assistance and current robotic systems (inspired by [Taylor 2006]), the applicability differs. To gain the necessary acceptance of robotic assistance, it is crucial

to exploit the strengths of robotic systems to complement man power while minimizing human flaws. In cooperation with humans sharing the same workspace, the risk exists that humans might come to harm. As stated above, the robot has to obey commands given by humans except such an order entails an expected danger for humans. For safety reasons, human operators must always be capable of interfering with the robot control [Albus 1989]. Orders can be delivered by using any kind of communication media and equipment. Furthermore, the robot has to protect its own existence as fare as possible for as long as it neither harms humans (strong requirement) nor assumes commands (weak requirement) that might risk life. However, the second condition is difficult to put to practice, as it is difficult to evaluate any potential risks.

Table 2.1: Properties of human and robotic assistance [Taylor 2006]

	Human	**Robot**
Judgment	Excellent	Poor
Care	Fatigue, inattention	Untiring and stable
Adaptability	Easily trained Versatile & improvisation	Hard to adapt to new situations
Information processing	Capability for integration and interpretation of multiple information sources	Limited ability to interpret complex information Integration of multiple sensory & numerical data sources
Nature	Hard to keep sterile	Immune to radiation and ionization
Hand-eye coordination	Excellent (human scale)	Limited
Dexterity	Excellent	Limited
Accuracy and strength	Limits in fine motion Bulky end-effectors (hands)	Excellent geometric accuracy Different scales in motion and payload

2.2 Technical requirements on service robots

Dealing with intelligent behavior requires a closer look on what intelligent behavior can be and how intelligent robots should be like. This requires a physical description as well as a definition of the required performance. The first characteristic has implications for the second, as it is directly coupled to the application and terrain the robot is dedicated to. The general focus of an intelligent machine like a service robot is to assist humans in any kind of context. For robotic systems for rehabilitation and assistance in daily life, the cooperation with humans is inevitable and the robot has to prevent injury to humans at any costs.

According to the Definition 3.14 of ISO, robots have to be specialized and programmable moving devices. Within this abstract definition, an enormous variety of robots of diverse shapes, all acting in various kinds of environments and equipped with different sensors and actuators, can be found. A robot has to be described by its physical structure, the so-called appearance, and its performance, founded by its behavior. Autonomous robots in domestic use have to fulfill requirements in both areas. A major difference between industry robots and the majority of service robots is that a service robot has to face the constantly fluctuating dynamic properties of its environment. Industrial robots are part of a predefined work flow in an environment that is attuned to the presence of robots. The main challenges in the design of service robots are the wide range of desired variability in an uncertain environment that is not adapted to robots. The uncertainty of the environment, in which the robot will be used, requires a higher degree of flexibility in the adaptation of the robot. Its physical functionality and

its behavior have to be robust enough for environmental changes and a variety of different service tasks.

2.3   Requirements on the physical appearance of service robots

Robots can vary in their physical shape in terms of size, material, type and number of joints, sensing system, locomotion system, but also in the potential of their computational system [Arkin 1998, p. 5]. Besides social and economic requirements, technological boundaries are determined by application-dependent conditions. In addition, the mechanical design has to consider aspects of the system like applicability and work space. Applying the pre-conditions of the previous chapter to the application of service robots in domestic area, the service robot has to fulfill the following basic requirements on physic and behavioral abilities:

−   The robot has to be capable of moving in the operating environment without external help. This entails the limitation of the physical structure of the robot, i.e. it must not exceed the dimensions of an average human body in order to allow for the necessary freedom of movement in domestic areas [IFR 2006, p. 416]. Nevertheless, the robot does not need to copy the appearance and size of a human body one-to-one. Furthermore, the robot must be capable of overcoming all obstacles that a healthy human can overcome without difficulties, which implicates similar functional and physical abilities like a human, e.g. open doors, walking up and down the stairs (if there are no elevators available), etc.

−   Domestic areas contain many obstacles such as furnished rooms [IFR 2006, p. 416]. Therefore, the robot has to be able to move under adverse conditions, e.g. narrow places, obstacles like stairs, barriers, furniture, moving objects (pets, humans, other robots or machines).

−   Compared to industrial counterparts, service robots shall provide lower acceleration and speed than heavy-duty robots, thus reducing the risk of dangerous situations (collisions) with humans [Bolmsjö 1995].

−   In order to ensure safety, the robot must be capable of observing its work space, which is shared with human operators [Schraft 2004, p. 30]. Due to its mobility, the robot needs necessary information to by-pass obstacles while maintaining task execution [IFR 2004, p. 416]. Therefore, a robot needs sensors which either facilitate it with perceiving its environment or enable it to use this data for low-level control.

−   The robot has to be capable of providing data of his current state in order to improve autonomous maintenance without external help.

−   The robot must be able to grasp and move objects of varying size and weight. Compared to industrial robots, the payload can be low, typically in the 5kg- range [Bolmsjö 1995]. However, the payload/weight ratio might be relatively high when compared with the size of the robot [Bolmsjö 1995].

Just as the robot needs a proper locomotion system for mobility, it will need an appropriate sensory system to operate under the given circumstances. Besides these general demands, the robot might face additional requirements, which are specific to its environment and application. In latter sections of this thesis, the physical appearance will not be discussed in detail except for the sensory facilities, which

build the basis for the perception of the environment and form a basic set of fundamental actions of the robot, which are necessary for selecting an appropriate behavior in task allocation.

2.4 Requirements on the behavior of service robots

Definition 2.1 of Roland Arkin: "Intelligence endows a system (biological or otherwise) with the ability to improve likelihood of survival within real world and where appropriate to complete or cooperate successfully with other agents to do so."[Arkin 1998, p. 31]

Behavior control has become a big issue in recent years [Bryson 2007]. From the technological point of view, the human mind possesses the following genuine capabilities, which are highly desirable for control systems [Ratanaswasd 2005] (comparison with Table 2.1 of [Taylor 2006]):

− Generation and monitoring of processes

− Filter and focus of task-related information

− Maintenance and adaptation of goal-relevant information

− Inhibition

− Different levels of abstraction from routine actions to complex deliberation

− Learning of new behavior in novel situations

The new model shall improve automation control systems and mobile robots to overcome the typical problems faced by classic Artificial Intelligence, e.g. the focus of attention, which is necessary to reduce perceived and evaluated data to the essential information that is necessary for the current task [Lueg 1997] and [Ratanaswasd 2005]. To act appropriately by fulfilling different services, the robot has to feature a fundamental understanding of the operating environment [Schraft 2004, p. 30]. The patterns of the service robot have to be adapted autonomously to the changes in the environment. Facing multi-robot systems, autonomous robots have to be capable of fulfilling a mission without additional help of a central coordinator [Jones 2005]. In this research, the following fundamental requirements have been assigned as crucial for robotic design:

− *Recognition of a situation*: using its sensor system, the robot shall not only be capable of collecting environmental and internal data but also of symbolizing these incoming data to give an image, a snapshot of the current environment as an instant (referred to the term "situation" in [McCarthy 2000, p. 9]) in combination with the internal state of the robotic system, which is set up to compare it with stored situations for evaluation. Therefore, sensory data has to be filtered and processed so the system can recognize elements and events in the external world.

− *Evaluation of a new situation*: the robot shall be capable of distinguishing between perceived images, evaluating them and labeling them with a general meaning as it depends on the temporary intention and service tasks in order to reduce the amount of appropriate actions and efficient action patterns [Ratanaswasd 2005].

− *Autonomous selection of actions according to the situation*: the main difficulty is to define which behavior is appropriate in the current situation. The proposed action might be conflicting in a long-term goal. To choose actions, it is necessary to predict the implications an action might cause

[Jones 2005]. In order to evaluate situations and make predictions, a fundamental idea of the environment, the world model describing fundamental relationships of the external world is necessary.

- *Flexibility in mission completing*: tasks can be similar and are categorized due to their main events and objects, but a task will almost never be identical with another task, as the positioning of objects, the time (daylight, night), the nature of objects (form and color of objects) may vary. Therefore, an adaptation to the specific, recognized situation will be inevitable [Ratanaswasd 2005].

- *Task decomposition*: similar to the approach of NASREM [Albus 1989], there are different levels of complexity to achieve high-level goals by splitting them into sequences of low-level actions.

- *Global memory*: in order to evaluate situations, predict the consequences of actions, different types of memories that contain data that might be perceived, evaluated, learned or initially inserted are necessary [Dodd 2005a] and [Ratanaswasd 2005].

The target of this project is to design and build a functional model describing internal mechanisms for autonomous behavior arbitration that enables a mobile robot to fulfill these requirements. The main goal of this model is to create a functional model for autonomous robot behavior that facilitates the robot to act properly according to its assigned mission and solve constraints due to the completion of the mission, without causing damage to the external world or harm to its own systems.

2.5   Requirements on the control architecture

Although the focus is on a special application, the synthesis with a common architecture assists in the exchange of algorithms for a technology transfer. As it is expected that the hardware is generally subject to changes [Orebäck 2003], the portability of a design is highly desirable. A platform-independent design allows shared software infrastructure and avoids effort duplication [Cote 2006]. Therefore, it is essential to define a design of higher abstraction considering following basic requirements of the underlying system [Orebäck 2003]:

- Robot hardware abstraction allows the implementation on different hardware solutions using various types of sensors and actuators without any additional adaptation of the general mechanisms. The abstraction can be derived on different levels and has to be balanced with efficiency [Orebäck 2003] and [Blank 2004]

- The extendibility facilitates the enhancement of additional software modules and hardware for new behaviors: The scalability on the architectural level helps to avoid bottlenecks and constraints [Orebäck 2003].

- The run-time environment shall be optimized to minimize overhead [Orebäck 2003].

- The reliability of tools and methods are key components of the architecture [Cote 2006]. In a multidisciplinary development, the repeat using of available solutions can only be carried out with universal concepts, as different theoretical background entails different requirements and optimization.

Based on the formal theoretical background of an abstract architecture, its implementation and configuration has to be discussed to allow accurate judgments about the capability and applicability. In order to meet the requirement of portability, the integration of a wide range of existing communication protocols and robotic standards seems to be highly desirable [Cote 2006]. This helps to emphasize a

great number of ideas and application areas. To achieve these objectives, the following overall characteristics of the software system shall be complied with [Orebäck 2003]:

− The design has to be simple in its implementation and interfaces.

− The design has to appear correct in all observable terms.

− The design has to be consistent.

− The design has to cover all important situations.

In contrast to experiments in real-world environments, the simulation system provides tools that allow an insight into the control process. This is a great benefit for the development. The simplifications involved with the simulation, however, shall not mask design goals, like robot specific functionalities, available volume, technology, power requirements, etc. that are crucial in the design of a concrete application with a specific type of robot [Mondada 1994].

3 Basic assumptions and definitions

For the research in this thesis, two main groups of disciplines of cognitive science, including psycho-dynamics, and engineering disciplines sharing the same subject, the mental lives of the human arche-type, which is seen in both groups from a completely different point of view, have been employed. In an interdisciplinary work like this it is especially important to define the terminology used to determine aspects that might be termed differently in the various disciplines. This chapter shall give a short introduction into the key terms used in this approach.

3.1 Assumptions in cognitive science and adaptation for technology

Definition 3.1: Cognition is exhibited through the characteristics of "short and long-term memory, categorizing and conceptualizing, reasoning, planning, problem solving, learning [and] creating" [Dodd 2005a]

R. Plutchik stated that cognition is "…the activity of knowing, learning, and thinking, of which emotion is a part… evolved over millions of years" [Plutchik 2001], which has been transferred by W. Dodd into technical terms [Dodd 2005a]. Although the proposed fundamental capabilities summarized in cognition as proposed in Definition 3.1 cannot be applied generally, it gives a good starting point that shows the key functionalities required for cognitive processing. In numerous research articles describing functionalities of the human mind, there is a logical separation made between the nervous system and the rest of the body, although the nervous system and associated brain are intrinsically linked to the body, integrated in biochemical and neuronal control circles. Both body and brain form the organism, whose interaction with the environment cannot be conducted by brain or body alone [Damasio 1994, p. 88].

In order to emphasize the information processing functions, which are in cognitive science generally labeled as mental processes, the main focus is on the brain. In general, actions are caused by commands of the brain after intermediate processing. However, the level of complexity of processing can vary. For example reflexes fall into the category of spontaneous rule-based responses, which initiate simple actions and crude behaviors activated by a certain external stimuli without any deliberation (These mechanisms have been emphasized in Freud's stimuli-response-theory [Freud 1915]). But in the human being, numerous brain structures are located between "input" and "output" sectors constructing and manipulating a steady stream of images organizing and categorizing concepts for interpretation, whose strategies can be acquired and the decision about selecting an appropriate response is made available. According to [Minskey 1986, p. 18], these strategies can be referred to as concept of agents in technology (see Chapter 3.3.2 for further details).

3.1.1 Image

According to A. Damasio, an essential demand for the mind is the ability to create images and process them internally in a process he called "thought" [Damasio 1994, p. 89]. All factual knowledge is processed in form of images, which is essential for reasoning and decision making. These images can not only be visual but also contain of sounds ("sound image"), olfactory or of other characteristics. In general, these momentary constructions of the brain can be divided into two categories:

- Perceptual images are generated due to sensing and evaluation processes reflecting the concrete changes of the environment and the organism itself [Solms 2002, p. 90], forming various images of different sensory modalities.

- Mental images are substantial replications of patterns that have been experienced at least once before. They are recalled from real past experiences and adapted to future plans [Damasio 1994, p. 97].

This conforms to the first definition of successful research by G. Pratl, who has defined "the inner representation of the outer word", as an "image of the real world" [Pratl 2006, p. 22]. However, it is important to emphasize that images do not entirely represent copies or miniatures. They are approximations of a varying accuracy formed by a complex neural machinery of perception, memory, and reasoning. Instead of facsimile pictures of events, things or words, images are rather an interpretation, reconstructing the essence and giving it some meaning in order to make it comparable. Based on these findings, a working definition for an image is proposed as follows:

Definition 3.2 (adapted for a technical system): A perceptual image has been formed by a perception system due to the sensory input via the sensory architecture. The image represents a snapshot of the environment and/or internal state of the organism or technological system and can contain other, simpler images. On the lowest abstraction level, images contain symbols that have been predefined and are grouped in context with abstract images. A perceptual image represents and contains current entities of abstract situations (images).

Definition 3.3 (adapted for a technical system): An abstract image is a mentally created template for a group of perceptual images containing the same essential symbols. An abstract image is the interpretation of a perceptual image and can be used for comparisons in reasoning. Images do not contain any time dependences.

The system can generate numerous images, image streams containing semantic constructions describing events, objects, etc. based on the symbols of the perception system. Images can be of different contents:

- Physical/physiological type: this image can represent objects, actors, movement levels, etc. perceived in the environment (the outer world) and physical inner state, e.g. the position, the energy level.

- Psychological type: a psychological image describes the inner states of the mind, which are often a result of interpretations of situations or physiological processes. This image possesses an emotion vector of primary (basic) emotions evaluating the image. It is an interpretation of the meaning of the image.

Analogous to the terms "scenario" and "scenario image" used by G. Pratl, which implies that a situation cannot be known completely, the image represents only some essential facts of a situation. In this approach, an image is the basic theme-element of the perception & recognition process: images are used as triggers to change internal states or behavior as they represent the essence of all data that is perceived and evaluated in one single moment. The image includes both worlds: the inner world (refer to Chapter 6.2.1) perception (homeostasis) as well as the outer world perception (environment). Due to the projection problem of Artificial Intelligence (AI), the determination of necessary facts for future

situations [McCarthy 2000, p. 10], which are derived by underlying symbolization methods, is imperative.

3.1.2 State and situation

Definition 3.4 of Antonio Damasio: A state is "an artificial, momentary slice of life, indicting what was going on in the various organs of the vast organism during the time window defined by the camera's shutterspeed" [Damasio 1994, p. 87]

Definition 3.5 of Antonio Damasio: "Homeostasis refers to the coordinated and largely automated psychological reactions required to maintain steady internal states in a living organism." [Damasio 1999, p. 39]

Definition 3.4 and *Definition 3.5* give a very good example of the cognitive view of the state of a biological system that is also applicable to mental systems. The main focus lies here on the time aspect: the state is an instant summary of conditions of a biological system and its environment, representing the inner world, which in organisms is the homeostasis. As neither a biological system nor its environment stands still, the state can only be seen as abstract, quasi artificially frozen. This is a major difference from the silicon counterparts: a state in a technological system can be achieved more easily and can hold for a time period far longer than a moment converging to zero.

Neuroscience, as well as physiologically founded concepts, differentiates between "channel"-based and "state"-based brain functions [Solms 2002, p. 33]. The information derived from the external world is mainly processed by channel-dependent functions, while the organism itself is organized in state-dependent functionalities. Furthermore, some theories make another distinction and define two types of information as theoretically different types of states: the "body states" and "mind states" of the continuously changing organism [Damasio 1994, p. 87]. The body states are mainly connected with the functions of homeostasis (Definition 3.6), while the mind states give the momentary logical condition of processes. In [Damasio 1999, p. 37] A. Damasio has distinguished three stages of processing concerning mind states, sorted by different levels of consciousness (Definition 3.12).

− State of emotion (Definition 3.9 and Definition 3.10) is triggered unconsciously by bodily states.

− State of feeling (Definition 3.11), containing besides physiological processes also entirely mental processes, are represented unconsciously.

− The conscious representation of the state of feelings is the conscious state in which the organisms, feelings and emotions are known.

In the proposed approach, the main focus of the control architecture is on the emotional states. Adapting the biological *Definition 3.4*, a state in technical terms represents "a situation during which some invariant condition holds" [Rupp 2005, p. 343]. Although "strictly speaking there are no such machines"as all changes and movement take place continuously, in technology there are many types of machines which can "be profitable thought of as 'discrete-state machines'"[Turing 1950]. In a time-discrete concept, each state has to differ sufficiently in its characteristics (values, objects, components, sub-states) from other states that "these states are sufficiently different for the possibility of confusion between them to be ignored" [Turing 1950]. Referring to theories about state machines, in electronics the state of a system has to contain all information required to calculate responses to all actual and

(near) future inputs without using historic inputs or outputs [eBahill 2006]. Both, biological and technological definitions give a similar understanding of a state concept, although their system properties are based on different hardware structures and processing methods.

Human behavior is based on mental processes, e.g. emotional operating systems, which have state-dependent functionalities rather than channel-dependent ones. They represent the changes in the body, which are observed by somatic monitoring structures [Solms 2002, p. 106]. This type of information processing is necessary for technical systems emulating the humanlike behavior.

Finally, to complete the terminology, a differentiation of the term "situation" as it is often used in diverse scientific disciplines, e.g. psychoanalysis, AI, but also colloquial language shall be provided. The terms "situation" and "state", but also "scenario" are used interchangeably, which might lead to misunderstandings. I want to emphasize here the terminology and interpretation used in my approach in order to make a clear distinction. [Ueno 1999] uses the term "situation" to describe a highly abstract *state*, which is extracted for cognitive learning processes. I will not employ this definition in this approach here, as according to previous definitions a state provides an inner representation of the temporary conditions of an organism or technical system. By contrast, a situation gives the general, time-dependent setting of the whole (real) world, in which the system is situated in. A situation cannot be captured in its entire fullness, but is partly perceivable (by the organism, or an external entity), restricting the information to a set of most valuable information sufficiently characterizing the current situation (compare with Chapter 3.1.2).This data entity that describes a situation, which is called the image (*Definition 3.3* and *Definition 3.7*) of a situation in this research, depends highly on the application of the agent (Definition 3.30 and Definition 3.32) and on the capabilities of sensing. According to [McCarthy 2000] formal theories in AI, an a-priori decision determines which phenomena and rules will be taken into account and form a so called "bounded informatic situation". In this research, the term "situation", as it is labeled and defined by other disciplines like AI (compare with Chapter 3.1.2), shall be commonly avoided. The term situation implicates a certain time dependency, but without the clear distinction as has been made for image (*Definition 3.3* and *Definition 3.7*) and episode (*Definition 3.6*). Considering the time dependency a situation inherits in a limited time frame, this can be split into a sequence of images that can be perceived by the organism. This gives a related meaning to "scenario" (details in Chapter 3.1.3), but the time line of a situation is comparatively short, just long enough to allow for an interpretation of the main temporary conditions and perceivable states of the world, considering also time delays while capturing sensory data.

3.1.3 Episode and scenario

A perceptual image, as defined previously, can in itself not give direct information about changes in the environment. As there is a continuous stream of images, a coherent sequence of images which can show observable changes and events in the environment, as well as (partly not observable) internal changes within the organism or technical system is the basis for decision making. In cognitive science, the term "episode" is not explicitly defined, although the term "episodic memory" can be found in several approaches, e.g. [Solms 2002, p. 99] and [Shadbolt 2003], and has been taken over in technical approaches like [Buller 2002] and [Dodd 2005a]. The following definitions shall be valid within this research:

Definition 3.6 (adapted for a technical system): An episode is based on the sequence of an endless number, but at least two, of different perceptual images. It describes the changes (differences) between these two abstract images on a coherent time line. It represents the entirety of an abstract episode.

Definition 3.7 (adapted for a technical system): An abstract episode is a stored template of a group of similar episodes. It defines a sequence of abstract states caused by abstract images. Abstract images can be part of more than one abstract episode, which gives a unique direction for transition.

This entails that two different abstract episodes containing the same pair of states can exist. An abstract episode describes exclusively at least one transition between two abstract images, called image transition. Due to complexity, the maximum number of 4^{10} transitions is preferable. Episodes of higher complexity shall be split into smaller sub episodes. The pair image and image' can define the transition between two states completely.

$$\left(State_A \xrightarrow{\quad image \quad} State_B\right) \neq \left(State_B \xrightarrow{\quad image' \quad} State_A\right)$$

The term "episode" has been chosen for this approach because M. Solms uses the term "episodic memory" for "memories of previous instances of the self in relation to object" [Solms 2002, p. 99]. M. Solms splits memories into "semantic" and "procedural" types to emphasize different functionalities to memorize which are necessary for the mind to be conscious over time. This terminology breaks with terminology used in earlier research [Russ 2003] and [Pratl 2006], which used the term "scenario" instead of "episode" to describe a corollary of events as a sequence of snapshots (images). In [Pratl 2006, p. 55], a scenario is a sequence of events identified by the system. The sequence is not limited to short intervals as in situations, but it can span a longer period of time. This is equivalent to the definition of "episode" in this approach. It is important that "episodes" are memorized (given or perceived in the past) and used for system control. In this approach, a "scenario" is not a control term, but it gives the sequence of events and situations in the environment, or in the case of a simulation it is an expected and prepared sequence of settings of the environment and robotic body to provoke situations for behavioral tests of the system.

3.1.4 Drive and stimuli

Definition 3.8 by Sigmund Freud: A drive is "the psychical representative of the stimuli originating within the organism and reaching the mind, as a measure of the demand made upon the mind for work in consequence of its connection with the body" [Freud 1915].

This well known term of drive ("Trieb")[11], as introduced by S. Freud in his metapsychological theories [Freud 1915] leads to a very controversial interpretation in cognitive sciences and psychoanalysis. S. Freud gives a comprehensive relationship between stimuli, drive and reflexes in [Freud 1915]. Freud describes the primary function of the nervous system in biological nature as an apparatus for

---

[10] This number was set arbitrarily to find a balance between two long and complex action sequences, which might take too long and are prone to interruptions, and two simple sequences, which do not have enough entropy to be valuable.

[11] The German term "Trieb" has no unified translation. It has been often translated in English literature with the term "drive", but the term "instinct" is also frequently used.

mastering stimuli, namely to abolish discharging stimuli with actions directed to the outer world, basic principle of the so called stimuli-response-schema [Freud 1915]. Many simple organisms are capable of performing actions spontaneously or reactively, responding to stimuli of two types [Damasio 1994, p. 89]: stimuli directed towards the outside and those of its origin within the organism itself.

For Freud a drive is a form of stimuli of psyche, which has its origin in the organism itself (homeostasis) and *cannot* be satisfied by simple reflexes [Freud 1915]. Although used explicitly by S. Freud, the term "drive" has diffuse transition to the terminology of emotions (Definition 3.9 and Definition 3.10), as the differentiation between drives, emotions and feeling is not unified. There is often a thematic intersection with the term "instinct." In [Solms 2002, p. 28], "drives" are basic motivations, which can be experienced as emotions. In [Panksepp s1998, p. 168] J. Panksepp even introduces a concept of "need states" opposing to the terminology of Freud in [Freud 1915], emphasizing that drives are generated by the bodily need detection system that rises the bodily need state to indicated regulatory imbalances, e.g. energy depletion [Panksepp s1998, p. 168]. This is congruent with the descriptions of a so called "drive state" of A. Damasio in [Damasio 1999, p. 77] and [Damasio 1994, p.] Drives (instincts) shall ensure the organism's survival, which is dependable on the complexity of the organism and the unpredictability of its environment [Damasio 1994, p. 123]. An important aspect was also added in the theory of A.Damasio, as he describes the direct connection of images with drives describing a "pervasive impairment of the drive with which mental images" [Damasio 1994, p.73].This goes along with the Theory of S. Freud in Definition 3.8. Drives are the mental representation of bodily founded stimuli. S. Freud donates explicitly that drives are the interface between the somatic and psychical processes, and their characteristics are arbitrary [Freud 1915].In this context S. Freud defines the following related terms that are necessary to complete the theory of drives [Freud 1915]:

− *Urge*: all drives appear urgent. The urge is the sum of the bodily forces giving the magnitude of the current demand.

− *Goal/ desire*: the general goal is the satisfaction of the drive, which can be achieved by abolition of the source causing the current drive. Although this final goal is unchangeable, there are various methods resulting to this common goal.

− *Source/ cause*: the cause of a drive is founded in somatic processing in the body and is not directly observable, but the source of a drive can be expressed by the emerging desires. The systematic study of the causes are not part of psychoanalysis anymore, but can be found in neuro-scientific contributions, e.g. [Panksepp 1998].

− *Desire object*: the most variable object of a drive is its desire object. It represents a tool, to satisfy a desire. There is no fixed intrinsic component relating desire objects and drives. They have to be assigned due to experience and can change through life time. A desire object can be part of the body itself, or extrinsic. One object can serve several desires.

According to S. Freud, all drives are congeneric, and can be distinguished solely by their causes. S. Freud restrains from defining a fixed group of drives, but proposes a general classification of two major types of basic drives: sexual drives and ego drives (self-preservation), which can be source of the psychic conflicts [Freud 1915]. Based on neurological findings J. Panksepp distinguishes four main drive-specific systems:

− *Hunger*: caused by energy depletion.

- *Thirst*: caused by water depletion.

- *Thermal balance*: caused by temperature changes beyond limits of optimal operation.

- *Sexual arousal*: caused by harmonic cycles.

The imbalance in homeostasis leads to a motivational state or drive state [Damasio 1999, p. 77]. There can be found some controversial statements within the context of drives. But in general it is widely accepted that drives are subconscious, evolutionary developed and relatively inflexible mechanisms leading to stereotyped motor patterns. The basis of instinctual behavior is closely linked to the inner balance of the organisms. Their goal is the survival of the organism. The basic idea of drives is to activate action systems or manipulate decision making to gain from the environment all necessities to sustain and recover from the imbalance. These systems act anticipatorily, activating regulatory behaviors. By activating or inhibiting emotional behavior systems, drives shall *prevent* future imbalances rather than correct existing ones.

As these drives compound with the special needs for sustaining bodies that are made of organic materials, these methods cannot be applied one-to-one to robotic physics. Nevertheless, as an embodied system the robot may face drives, which are distinct to those of an organism. These aspects shall be discussed in Chapter 6.3.1.

3.1.5 Emotion

Definition 3.9 by Mark Solms: "Emotion is akin to a sensory modality – internally directed sensory modality that provides information about the current state of the body itself, as opposed to the state of the object world" [Solms 2002, p. 105].

Definition 3.10 by Antonio Damasio: "Emotion is the combination of a mental evaluative process, simple or complex, with dispositional responses to that process, mostly towards the body proper"… (called emotional body state)…", but also towards the brain itself, resulting additional mental changes" [Damasio 1994, p. 139].

According to *Definition 3.9* emotions seem to have their roots in physiological processes and the perception of them, but, in contrast to drives they are situated on a higher level of abstraction in relation to the environmental situation. Emotionally based behavior is more flexible and sophisticated than the drive-based one. The methodology for classification of emotions seems to be rather unclear in cognitive science and psychoanalysis. But many approaches agree that there is a group of primary emotions that is directly linked to the innate functions leading to overt behavior, which is evolutionary necessary to adapt to environmental changes. These behavior patterns are mainly stereotyped motor patterns that are the basis for so-called instinctual behaviors and emotional expressions [Solms 2002, p. 29]

Primary (basic, universal) emotions

The number and nature of emotions is a widely discussed topic in cognition science leading to numerous controversies. Damasio [Damasio 1999, p. 50], Plutchik [Meyer 1999, p. 151] or McDougall [Meyer 1999, p. 113] have introduced a group of basic or primary emotions, which vary in number and denotation. Unlike these approaches, Panksepp introduced emotional control systems, so-called

"affect programs" [Panksepp 1998, p. 41] of the human brain that represent a limited group of basic emotional circuits emerging in universally recognized emotions. There are four primary emotional control systems, which have a high number of manifestations on the neuronal level besides external manifestations [Panksepp 1998, p. 51]:

− Seeking-System: provides positive incentives and can be labeled as "curiosity, interest". The foraging shall help to orientate oneself and find new positive stimuli. The seeking system is the only emotional operating system that has no pre-defined objective to be achieved. This system will be triggered by inner states, mainly drives which arise due to inner imbalances. The seeking system awakes the appetitive interest in the object world [Solms 2002, p. 211].

− Panic-System: can be found in the context of social bounding. It shall prevent social loss. The objective of this emotional control system is to find congeneric individuals to build up social bounds, which might improve the probability to survive. Actions of this control system can inhibit action pattern of the seek system in certain situations, e.g. the return to the group instead of a continuing exploration.

− Rage-System: shall provide mechanisms to protect against irritation, restraint, frustration, etc. The objective of this emotional system is to overcome obstacles and barriers. It is also necessary to satisfy elementary needs, which might come into conflict with the interests of other individuals. Potential actions of this control system are push, hit, destroy and fight.

− Fear-System: has its origin in pain and fear of destruction. Its objective is to prevent the individual from potential danger. The majority of actions aim at increasing the distance between the source of danger and the individual, e.g. escape, hide, etc.

The available theories approve the idea of primary (basic) emotions as the basis for secondary (complex) emotions, which is based on neurological experiments [Panksepp 1998, p. 50]. This concept shall be the archetype of control architecture in this approach.

Each emotional system interacts with numerous others in a highly dynamical way on diverse hierarchical levels of the brain in order to facilitate behavioral choices. However, emotions are far more than a rule-based filter channeling activities of the cognitions systems. They lead to spontaneous and either genetically caused or learned action responding to changes in the environment. They have an anticipatory character controlling a vast range of the cognition apparatus in order to anticipate possibilities preventing arising difficulties rather than reacting to those that actually do occur. The emotional evaluation is strongly linked to the cognitive apparatus [Panksepp 1998, p. 39], which makes it difficult to distinguish them from the underlying control mechanisms of the outward, observable behavior.

Secondary (complex, social) emotions

In addition to the basic mechanisms of primary (basic) emotions, there are further mechanisms of secondary emotions building the body of systematic rules and regulations, connecting "categories of objects and situations, on the one hand and primary emotions, on the other" [Damasio 1994, p. 134]. In contrast to the primary emotions, these rules are not congenitally pre-defined but socially formed during the organism's life time. The emotional taxonomy can be broadened by a far larger number of emotions of higher-order constructs during the individual development, which are in a context of terms like social emotions, feeling states and motivation. In general, these are also established in psy-

chological approaches [Solms 2002, p. 114]. Dealing with indirectly psychological and physiological reflections based on the dichotomous distinction of approach versus avoidance indicates distinct varieties of emotional behavior. A more detailed discussion of this topic will follow in Chapter 5.4.
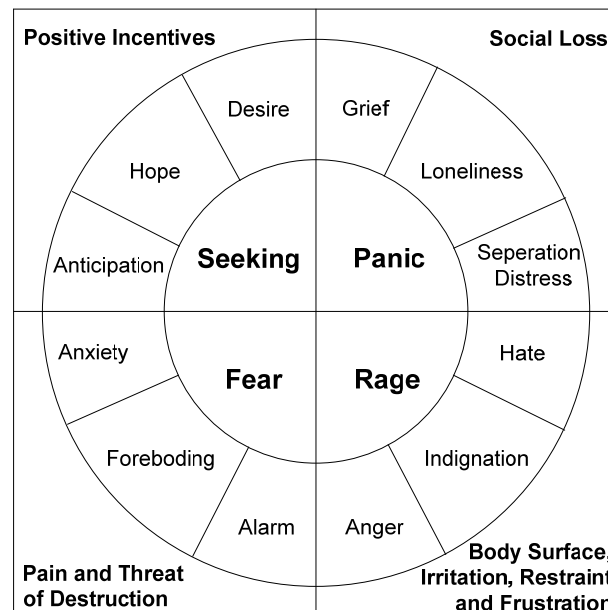
Figure 3.1: Primary emotional systems due to environmental influences adapted to [Panksepp 1998, p. 50]

Background emotions

This term has been introduced by Antonio Damasio to describe the subtle details of minimal changes in body posture, speed, contour movements. These minimal changes in the physical state provide the general tendencies, the impression that a person might be, e.g. "cheerful" or "tense" [Damasio 1999, p. 52]. This type of emotion is internally caused by the process of homeostasis, which is also the process of the satisfaction or inhibition of drives and motivations. The main difference from other forms of emotion is that it is not directly apparent and does not cause any observable action, but it might appear as an offset in the current emotion vector. This form of emotion is not directly considered in this approach, but it can be deployed by the remainder of former emotion vectors.

In any case, emotions are interpersonal concepts, which can be found in their basic form in all individuals. This is in contrast to the concept of feelings, which is directly linked to the personality and consciousness of the individual (Definition 3.12).

3.1.6 Consciousness, mind and feelings

Definition 3.11 by Mark Solms: "The perception of visceral information is registered consciously as feelings of emotion."[Solms 2002, p. 29].

Definition 3.12 by Antonio Damasio: "...the term feeling should be reserved for private, mental experience of an emotion, while the term emotion should be used to designate the collection of responses, many of which are publicly observable" [Damasio 1999, p. 42].

Definition 3.13 by Antonio Damasio: Consciousness is "an organism's awareness of its own self and surroundings" [Damasio 1999, p. 4].

In M. Solms theory, the terms "feeling" and "emotion" seem to be inseparably amalgamated (*Definition 3.11*). This stands in sharp contrast to *Definition 3.12* by A. Damasio. According to A. Damasio, there should be a sharp distinction between the term "feelings" reserved for the individual mental experience of emotions, which are directly linked to the consciousness (*Definition 3.13*) and "self" of an individual, while "emotions" are a more basic and fundamental and partly observable form of response. Based on this determination of feelings, which I will use for further descriptions, feelings are directly compounded with the consciousness of an individual. M. Solms stated in [Solms 2002, p. 28] that feelings, personal (autobiographic) memory and consciousness seem to be inextricable. A main challenge in cognitive science is to respond to the operation of consciousness, whereas the description is often given in quantitative rather than qualitative terms [Solms 2002, p. 86]. S. Freud stated that consciousness is just a limited part of the mind and the majority of mental functions operate unconsciously. This idea is also used in current cognitive science. The state of consciousness does not only reflect the concrete changes of the outer world, but also represents the internal changes of the organism. Conscious states create a virtual body and can be categorized in three global states: awake, aware and alert [Solms 2002, p. 85]. According to M. Solms [Solms 2002, p. 92], A. Damasio's neuroscientific view goes along with Freud's psychoanalytic theories in describing that the internal state is directly and at the same time linked to the state of the object world (outer world in Chapter 6.2.1). A. Damasio calls this coupling "core consciousness".

There are many theories and metaphors trying to explain the mechanisms of consciousness in terms of information and cognitive science partly trying to reformulate the psychoanalytic concepts [Baars 1996]. Corresponding to the philosopher G. Strawson, M Solms agrees that the essence of the mind "is not intelligent behavior but, rather, subjective awareness" [Solms 2002, p. 71]. A main concern in this context is to relate consciousness to the mind. M. Solms claims in [Solms 2002, p. 71] that analogue to S. Freud the consciousness is not synonymous with mind, but merely a property of the mind, introducing the idea of an unconscious mind possessing thoughts, intentions, memorized data, etc. of which a person is not consciously aware.

3.1.7   Memorizing and memory systems

Definition 3.14 by Mark Solms: "'working memory' is synonymous with the ability to consciously 'hold things in mind'"[Solms 2002, p. 83]

Definition 3.15: "Episodic memory is concerned with unique, concrete, personal experiences dated in the rememberer's past."[Tulving 1983, p. v]

Definition 3.16"Semantic memory refers to a person's abstract, timeless knowledge of the world that he shares with others" [Tulving 1983, p. v].

Researchers in artificial [Tulving 1983, p. 32], [Minsky 1988, p. 153], [Braddeley1997, p. 38], [Nuxoll 2004], and [Dodd 2005a] and natural intelligence [Solms 2002, p. 81], [Damasio 1999, p. 220], and [Sacks 1986, p. 103-110, p. 125-142] agree: the human mind possesses different types of memories. However, there is no agreement on a consistent categorization of these memories. The role of memory is crucial in both domains, influencing each other in the way they are looking at it [Turkle 1989]. In the following I will give a short introduction into the main hypothesis of memory in

psychoanalysis and cognitive science and compare it with technological approaches in computer science.

Looking back on a long-established history, there are numerous approaches in cognitive science that are concerned with this topic. They have been introduced and summarized in both [Tulving 1983] and [Braddeley 1997] in detail. In the human mind, the semantic memory showing causal connections, the episodic memory containing autobiographical memories and the working memory, whose capacity determines the ability to hold visual-spatial information in mind, play an important role [Solms 2002, p. 97]. The semantic memory holds abstract knowledge that permits an individual to work with inner representations in a goal-directed manner [Tulving 1983, p. v] and [Wagner 2002], and has been the basic concept for a long time. Furthermore, there are functional visuo-object and visuospatial representations recovering and evaluating the meaning of semantic control [Wagner 2002]. Besides abstract information storing, the episodic encoding transforms experience into a memory entry that can be consciously remembered [Tulving 1983, p. 41].

A. Damasio introduced the term of the "autobiographical memory" in [Damasio 1999, p. 217] and [Damsio 2003, p. 270] to emphasize the systematic memorizing of situations and multiple instances of stored individual experiences, showing the main biography of an individual. This type of memory storage is the sum of all memories, which have been gathered during the (mentally active) existence of the individual, and form the individual's specific history and experience. This memory grows continuously during a life time, but will be partly reorganized with new experiences. This memory type is based on the autobiographical self [Damasio 2003, p. 270]. First described by E. Tulving in [Tulving 1983, p. 28], this type of procedural memorizing that forms an individual autobiography is a functionality of the episodic memory, which has been ignored by computer science for a long time [Nuxoll 2004]. In recent approaches, e.g. [Nuxoll 2004] and [Dodd 2005a], episodic memory has become an essential part in cognitive control.

Generally, the existence of a short-term memory (STM) and a long-term memory (LTM) has been proved in cognitive sciences [Braddeley 1997, p. 38] and [Tulving 1983, p. 23]. While the short-term memory holds copies of current entries originating from perception, these are compared with entries in the long-term memory [Tulving 1983, p. 23] and [Dodd 2005a], and in the case of episodic memory with knowledge of abstract episodes (*Definition 3.7*). In some technical approaches [Braddeley 1997], [Buller 2002] and [Ratanaswasd 2005][12], a third type of memory has been modeled: the working memory. The working memory represents an alliance of temporary memory systems, containing information on the current environment, and relevant knowledge [Ratanaswasd 2005]. This concept is congruent with the idea of [Solms 2002, p. 97] and shall be a guiding principle in this approach.

Congruent with conscious and unconscious perception there is also conscious and unconscious memorizing [Wagner 2002] and [Solms 2002, p. 81]. An implicitly memorized entry as complement of explicitly memorized entries cannot be accessed deliberately as explicit memory entry, but it can initiate diffuse tendencies, which might have an effect if there does no specific memory entry exist (e.g. in case of a new situation) [Solms 2002, p. 81]. This can be seen in analogy with data management functionalities. Additionally, [Braddeley 1997, p. 32] proposed that old and rarely retrieved

---

[12] This approach refers to A. Damasio's theory.

entries of long-term memories will be made less and less accessible, which can be derived by sorting rules and expiry dates [Roesener 2007].

3.2 Definitions in robotics

Robotics is a research area that concentrates on the design, programming and control of an intelligent machine, humanoid or not, called robot (derived from the Czech word "robota": "compulsory labor" first used by K. Capek in 1921 [eCapek 2006]). More than 50 years ago, the first numerical machine was built on MIT, followed by the first programmable robot designed by G. Devoland and J. Engleberger in the 1950s-60s. According to R.C. Arkin [Arkin 1998, p. 5], robots are machines that can be programmed to fulfill a vast range of different missions. Besides the special terminology of mechanics and kinematics, the following chapter shall provide an overview of the general vocabulary used in robotics and robot control as it is used in this approach.

3.2.1 Industrial and service robot

Definition 3.17 by the International Standardization Organization: A manipulating industrial robot is "a re-programmable, multi-functional, manipulator designed to move material, parts, tools or specialized devices through variable programmed motions for the performance of a variety of tasks" [ISO/TR 8373: 1994, p. 2, Definition 2.6]

Definition 3.18 by the International Standardization Organization: A mobile robot is a robot (of Definition 3.14) which carries all of the means needed for its monitoring and movement (power, control, driving) [ISO/TR 8373 1994, p. 3, Definition 2.13]

Definition 3.19 by the International Standardization Organization: "robot system comprising
- robot (of Definition 3.17);

- end effector ([ISO/TR 8373 1994, p. 5, Definition 3.11]: *'device specifically designed for attachment to the mechanical interfac to enable the robot' (of Definition 3.17)' to perform its task');*

- any equipment, devices, or sensors required for the robot to perform its task;- any communication interface that is operating and monitoring the robot, equipment, or sensors, as far as these peripheral devices are supervised by the robot control system" [ISO/TR 8373 1994, p. 3, Definition 2.14].

Definition 3.20 by Roland Arkin: "An intelligent robot is a machine able to extract information from its environment and use knowledge about its world to move safely in a meaningful and purposive manner" [Arkin 1998, p. 2]

Definition 3.21: Service robots are "freely programmable automated or semi-automated mechanical devices designed to perform a service rather than a manufacturing function." [Schofield 1999]

Definition 3.17, Definition 3.18, and Definition 3.19 define major aspects of robotics. They were published in 1994 by the technical committee TC 184/SC 2 "robots and robotic devices" of International Standardization Organization, giving unified vocabulary as a part of the Standard for "manipulating industrial robots" (SO 8373:1994). Comparing the ISO definitions with Definition 3.20, R. Arkin gives a far more behaviorist view of the term "robot", which might fit better into functional

concepts as it implies the relation of the robot to its environment. Congruent with the statement of [Shell 2005] that "people and robots are embodied within and act on the physical world", J. McCarthy defines a robot as an embodied agent (for details see Chapter 3.3.2), which has the bodily and mental capabilities to perceive [13] its environment in order to move and interact within the environment [McCarthy 2000, p. 5]. In general, the term "robot" in Definition 3.20 will be used in this approach. Based on the assumption that mobility and autonomous moving are crucial requirements for this research area (Chapter 2.3), the term robot refers here always to a mobile robot (Definition 3.18), which possesses sensors and communication devices of the robotic system (Definition 3.18), allowing self-monitoring and mission completion without external devices. The robot's type and level of mobility is not defined, but the mobile robot is limited in our approach to earthbound movement with the capability of moving on planes with a different level of inclination and after overcoming minor obstacles. There is a vast range of robots of different shapes and functionalities. Due to its control system, the term "logical robot" was established in AI, describing a robot that is controlled by a program (on board or externally), which executes goal-oriented action selection based on logic concepts [McCarthy2000, p. 13].

The definition of service robots is not unified. However, Definition 3.21 gives a very good interpretation of this term. The main difference between Definition 3.17 and Definition 3.21 of industrial robots lies in the application of tasks. The applications can cover various domains, such as cleaning, observation, or undersea applications, which might entail different requirements. They have in common that the working space is remote, less predictable, unstructured, rapidly changing and unstructured, requiring a higher degree of autonomy, reliability, and adaptability in control. There are similar technologies considered to have the potential to resolve this by forming an independent research area [Shofield 1999].

3.2.2   Control system and robotic control

Definition 3.22 by the International Standardization Organization: A control system is a "set of logic control and power functions which allows to monitor and control the mechanical structure of the robot and to communicate with the environment (equipment and users)" [ISO/TR 8373: 1994, p. 2, definition 2.7].

Definition 3.23 by the International Standardization Organization: Adaptive control is a "control scheme whereby the control system parameters are adjusted from conditions detected during the process"[ISO/TR 8373: 1994, p. 1, definition 5.3.4].

Definition 3.24 by the International Standardization Organization: learning control is a "control scheme whereby the experience obtained during previous cycles (6.22) is automatically used to change control parameters and / or algorithms" [ISO/TR 8373: 1994, p. 1, definition 5.3.5].

Definition 3.25: Robotic control is „the process mapping a robot's sensory information into actions in the real world"[Jones 2005].

---

[13]The functionality of perception does not necessarily include the interpretation of the collected sensory data.

In our approach, the term "control system" as defined in Definition 3.22 contains a broader range of high-level control functionalities beyond the control of the robotic physics. There are different types of control systems, e.g. providing adaptive control, which updates in processes due to the environmental situation (Definition 3.23). A learning control system can use historic data for optimization (Definition 3.24). The control system contains software and hardware. In this approach, the focus is mainly on software architecture. In analogy to the human mind, where free choices are considered to be free will, a robot needs a system to consider the choices leading to the desired situation. In analogy to the human mind, which is capable of choosing between alternative actions, the autonomous control system of a robot can be considered as the robotic free will [McCarthy 2000, p. 7].

Definition 3.25 describes the general relationship between sensing and acting. Every robot is required to complete actions that influence the environment and are influenced by the environment. In the case of robotic control as it is used in Definition 3.25, the mapping of the sensory information can give a wide range of potential actions and behaviors. The choosing of a particular action from a group of potential behaviors is called action selection and is a major functionality in robotic control [Jones 2005]. Facing the whole spectrum of behaviors, the selection of actions and action patterns require reasoning, coordination, planning and fundamental knowledge about the inner constraints and rules of the (outer) world to avoid collisions, bottle necks and behavioral inconsistencies. Robots are controlled by using sets of interacting behaviors, containing actions and action patterns, in order to achieve desired goals. A classical approach is [Jones 2005]. Even in simple, reactive behavior, which represents an input-driven type of control, the rule-based stereotyped behaviors inherit (in this case predefined) knowledge about the interdependencies.

The range of capabilities of behavior-based control architectures is of a great diversity. Behavior-based control does not rely on a complete world model, maintaining the constraints of sensing and acting. There are different approaches controlling single robots and/or considering the coordination of a robotic team [Jones 2005].

3.2.3   Configuration

Definition 3.26: "The specification of components, connections and structure of the control system for a group of robots will be called configuration" [MacKenzie 1997].

In general it is important to distinguish between theoretical control architecture and its implementation [Orebäck 2003]. To test an implementation a collection of active components and their communication have to be specified forming the configuration of a specific control architecture. This will be a major topic in Chapter 2.5, as the configuration represents the missing link between the universal concept and application-specific architecture designed to fulfill a certain range of tasks. The configuration is also an object of optimization, and responsible for the performance of the architecture. The configuration, as requested in Chapter 2.5 aims to describe a single system and not directly emphasize a team structure, which might be the focus of multi-robot systems (described in the following Chapter 3.2.3.).

3.2.4 Multi-robot system (MRS)

Definition 3.27: "A multi-robot system (MRS) is here understood as a system which consists of several autonomous/semiautonomous robots working together trying to achieve a common goal" [Taipale 1993].

Definition 3.28: A multi-robot system "is a system composed of multiple, interacting robots"
[Jones 2005].

While Definition 3.28 defines the system property only in a very general manner, Definition 3.27 gives a more detailed interpretation of MRS as it emphasizes the autonomy of the robots and reasons that their cooperation is goal-oriented. In the context of the definition of an agent (Definition 3.30 in Chapter 3.3.2), the term multi-robot systems (MRS) can be used synonymously with multi-agent systems (MAS), which proved to have similar characteristics from a computer-scientific point of view. A major difference lies in the fact that robots as well as human beings are embodied and interact with the real world [Shell 2005].

Although it makes sense that cooperation is goal- oriented, the independent operation of robots is not crucial for the existence of an MRS. In centralized MRS and centralized MRS control the actions of robots are determined partly or entirely by external commands given by an outside entity, which might be a main frame or even other robots [Jones 2005]. The main focus in our research is on distributed MRS, where robots have to make their own control decisions based on their limited, local sensory information [Jones 2005]. A cooperation with other systems, like control networks, is not excluded in the latter form of MRS, but the main difference lies in the fact that there are no external entities involved in the accomplishment of tasks assigned to the MRS. The advantage of this type of MRS lies in the functionalities of scalability and performance in uncertain and unstructured environments, as local data acquisition and decision making can compensate for these uncertainties by giving a more flexible, fault-tolerant system that is built up by comparing simpler robotic entities.

In general, robotic teams can be characterized by following properties [Arkin 1998, p. 369] and [Panait 2005]:

− Team size

− Communication range

− Communication topology

− Communication bandwidth

− Team composition and reconfiguration

− Individual processing ability

These properties define domains of the team characteristics as well as the capabilities of the team entities [Sugawara 1999]. The basic idea and benefit of multi-robot systems is to enable a group of simpler robots to achieve goals via cooperation, which even a sophisticated robot cannot accomplish on its own. The major benefits of MRS are:

- Greater error sustainability and fault tolerance: With the redundancy of robots, this system is less error prone, as the failure of one or more robots has less impact on the overall system.

- Task enablement and improved performance: allowing distributed sensing and parallel actions on different locations, which facilitates a far wider range of applications and capabilities.

Exploiting these beneficial properties, robot teams seem to be suitable for applications containing following tasks of parallel and distributed processing, e. g foraging, observation, formation cooperative transportation [Sugawara 1999]. However, these advantages are confronted with a number of problems. In these systems, there is a higher risk of robot-robot collisions [Arkin 1998]. Furthermore, cooperation requires communication, which sends up communication costs by requiring additional systems (Hardware and Software). Therefore, it is prudent to think of a team's application before designing robot teams. Tasks, which advance behavior like foraging, grazing or consuming, require the potential of a robotic team.

To organize robotic teams, the level of independence for robotic team members has to be carefully selected, which will have a direct influence on the forms of cooperation that might be actively coordinated and un-coordinated. This has also a high impact on the form and complexity of communication. Using different methods like motion planning or distributed AI shall help to overcome these control problems, giving the system self-organizing capabilities without further complexity. To ease the problematic constraints in a MRS-based approach, the majority of proposed systems focus strictly on homogenous teams, where robots are interchangeable and equally equipped.

3.3   Definitions and methods in Artificial Intelligence

The approach of Artificial Intelligence (AI) is to develop programs (algorithms) that allow computers to emphasize behavior, emulating humanlike capabilities indicating "intelligence" [Brooks 1991]. In robotics, AI provides concepts for robotic control. Conventional AI tries to build top-down systems [Brooks 1991] using designs based on the principle of "sense-model-plan-act" [Mochida 1995]. Although mainly associated with computer science, Artificial Intelligence is traditionally linked to other research fields such as mathematics, mechanics, but also psychodynamics (see Chapter 4.1) and other disciplines of cognitive science, and many others. This has led to a large number of various concepts, which have been established in AI research. Due to the analogy to this research, terms and definitions of related concepts shall be discussed.

3.3.1   Intelligence

Definition 3.29: „Intelligence endows a system (biological or otherwise) with the ability to improve its likelihood to survive within the real world and where appropriate to compete or cooperate successfully with other agents to do so."[Akin 1998, p. 31]

What is intelligence in the context of the human mind is a main issue, which has to be defined to determine the demands for its artificial counterpart. But according to H. Gardner [McMullen 2002], intelligence is not one property but a range of different skills which can be categorized. Intelligent behavior is indicated by its existence in biological systems. But where does intelligence start and where does it end? There is no general answer to this question, but it is considered that higher forms of

animals posses intelligence and express intelligent behavior. Animals seem the proof that intelligence can be achieved, although the necessary mechanisms are poorly understood [Arkin 1998, p. 31]. Human beings provide a far wider spectrum of intelligence in different forms of intelligence. These can be generally categorized according [McMullen 2002] to cognitive (IQ), emotional (EQ), and spiritual (SQ) types. In general, cognitive intelligence, containing e.g. mathematical and linguistic skills, logical deductions or rational thinking in general, is dominating the testing of human intelligence. But as B. McMullen stated in [McMullen 2002], "Being "clever" in the traditional sense is no longer enough". This form of intelligence cannot account solely for the overwhelming capabilities of the human mind in everyday life. Following H. Gardner's categorization of function-based intelligence, the emphasized spatial/ visual intelligence cannot be accounted for solely by mathematical, logical thinking. The spatial/ visual intelligence describes the capability of forming a visual representation of the environment in one's mind. This would be absolutely impossible with pure logical methods. The human mind has to cope with a vast amount of information, which has to be reduced and evaluated in recognition systems. Emotional intelligence, which can find a faster connection between internal states (and needs) and the sensed environment, plays a key role in this context. But it is exactly this capability that requires an inner representation of the world (Chapter 6.2.1), which is a key factor for AI. As emotional rather than cognitive knowledge gives a consistent inner representation of the world (Chapter 6.2.1), it shows that there is a theoretical flaw in classical research of AI.

3.3.2   Agent

Definition 3.30: "We define an agent as an active, persistent computational entity that can perceive reason about and act in its environment, and can communicate with other agents" [Singh 2003].

Definition 3.31: „An agent can represent knowledge of its world, its goals and the current situation by sentences in logic and decide what to do by interfering that a certain action or course of action is appropriate to achieve its goals" [McCarthy 2000, p. 2]

Definition 3.32: "An agent is a computational mechanism that exhibits a high degree of autonomy, performing actions in its environment based on information (sensors, feedback) received from the environment." [Panait 2005]

The term agent has a great number of different definitions in the scientific community, as the examples show above. The term agent is often used in a micro as well as in a macro sense. An agent can be either a small entity without any intelligence in itself or a more complex agent ranging up to the functionality of a whole organism, or of robot which in itself can be interpreted as an agent in the real world. Due to the "pseudo-mental oriented terminology" of [Shoham 1993], an agent can be any entity whose state (Definition 3.4) consists of mental components including e.g. choices and commitments. M. Minsky stated in [Minsky 1985, p. 18] that the mind is a schema of many smaller processes, each mindless in itself, which he called agents. The substantial interest of AI lies in building generally intelligent, autonomous multi-agent systems. Sentences in logic represent the knowledge of the agent's microcosm and accurate situation [McCarthy 2000, p. 1]. In our research, Definition 3.30 and Definition 3.32, which appear to be very similar, seem to be preferable as they are not bound by the sentences of logic (a very typical aspect in AI) as described in Definition 3.31, or implicated with mental abilities like intelligence, which might be mixed up with terms used for organisms. Due to

Definition 3.30, agents are not necessarily the smallest entities and may contain simpler agents, which can be split down again, and Definition 3.32 emphasizes the interaction of an agent in the environment, which goes along with the view of [Maes 1994]. Therefore, a computational entity can be called "agent" in this approach as long it possesses the following properties:

− Perception of its environment

− Deciding and execution of actions in its environment

The use of the terms "agent" and "robot" cannot be used interchangeably, but they are functionally connected. As we have seen in previous definitions, agents can be of all kinds of complexity, as long as they show the crucial characteristics as mentioned above. The main difference is that an agent is computationally and completely described and does not necessarily posses a physical body (i.e. is "embodied"). However, using Definition 3.20 for robots, an autonomous (not necessarily mobile) robot can be described as an *embodied agent* of higher complexity, as it is capable of acting and reacting in the (structured or unstructured) environment of the real world due to the perceived and collected environmental information. A robot is embodied as long as it has a physical body, which is limited by its physics, uncertainties and also unpredictable consequences of actions [Jones 2000]. A main difference between a computational agent and its embodied counterpart is that a robot is directly placed in the real world and acts according to the sensory information retrieved from the real world. [Jones 2000] describes this with "a robot is situated".

3.3.3   Common sense information and approximation

The knowledge derived from sensing and memory is crucial for the cognitive control system (see *Definition 3.22*) in order to make decisions about the next actions [Dodd 2005a]. However, as defined in formal theory, the set of relevant known facts are necessarily incomplete [McCarthy 2000, p. 4]. Formal approaches of AI often use the rule-based filtering of perception. However, these approaches lack humanlike abilities, as a machine has to decide autonomously which data are relevant and its reasoning cannot be derived monotonically [McCarthy 2000, p. 3]. Other than in formalistic approaches of AI, in which the relevance of events are predefined, non-monolithic reasoning uses common sense mechanisms that allow to evaluate dynamical situations without *a priori* limitations [McCarthy 2000, p. 4]. Key problems in this AI approach are the definition of the set of facts that the theory requires to change and the modification of behavior. Even in common sense information-based decision making, clear definitions are inevitable. The main characteristic of approximate control concepts is that although they are not well defined, their contents are defined. There are two assumptions regarding approximation [McCarthy 2000, p. 8]:

• It is considered that a precise concept exists but remains unknown because of incomplete knowledge.

• The concept is intrinsically based on approximation. In this case, the concept is not fully determined by nature and cannot be determined by human convention.

In our approach, the regulation will be derived from emotion evaluation combined with memorized episodes and desires as a dynamic filter of perception.

3.3.4 Reasoning

Due to AI, there are about three general types of reasoning mentioned:

- Causal reasoning concentrates on the consequences of actions and events, to determine laws to approximate the consequences of events [MyCarthy 2000, p. 9].

- Non-monotonic reasoning endeavors conclusions based on the "best" models of facts, which are continuously added to existing ones. As new facts might suspend current conclusions, the behavior shows significant changes through time [MyCarthy 2000, p. 15].

- Probabilistic reasoning is a special form of non-monotonic reasoning using probabilistic models to define a space of sample "events". However, this method is often used by humans for pre-definition rather than derived automatically by computers [MyCarthy 2000, p. 15].

All types of reason have as common, mutual basis that they are founded on logical, mathematical thinking. They are all part of cognitive intelligence (Chapter 3.3.1), which is not solely responsible for decision making and behavior arbitration in the human mind. This form of reasoning is often preferred in scientific circles, as it can be expressed easily with mathematical methods. Although very useful in its basic concepts, an entirely mathematically founded reasoning for robotic control shows flaws in the case of vast data amounts, which need a-priori filtering and evaluation, before they can be used for reasoning.

3.3.5 Time, continuous processes vs. discrete processes

In the human mind, a high number of continuous processes are event-oriented and use discrete approximation. As already mentioned in the definition of "image" (*Definition 3.3* and *Definition 3.7*) and "episode" (*Definition 3.6*), time plays an important role in biological systems as well as in every control system. Although not directly emphasized, the idea of timely and real-time processing is an important topic. But what is "real-time" in an organism? As already stated by I. Kant in [Kant 1781], time is part of the intellectual structure, where the human mind sequences events rather than measures time objectively. On the contrary, the "experienced time" of an organism is input-driven. Expecting an average amount of environmental and internal, mental changes, there is a "sense of time" in the mind, which makes the forming of a chronology in behavior and building of sequences in actions inevitable. However, a human does not have absolutely objective clockworks like a machine. The mental time in the mind is not as continuous as it might appear. As the processing is state-driven, the changes of states are mainly and significantly responsible for the experienced time. Furthermore, emotional states and strong physiological imbalances might also influence the "sense of time" dramatically.

As time seems a completely different value in mental processes, the question now arises which role it shall pay in the technical contemplate? Continuous processing might be neither necessary nor desirable. Although sensing processes will have to be continuously executed, the sequence of decision making might not require this continuity. Processing new steps shall be input-driven as well: only if there is a change of state shall behavior processes be adapted, new behavior selected and the scheduled actions eventually prioritized. Considering that only discrete processing is used in our approach entails the question about the continuity of actions, which shall be discussed in Chapter 6.3 in detail.

3.3.6 Learning

Although there are no clear boundaries, there is a general tendency to distinguish behavior between "built-in" and "learned" [Minsky 1988, p. 115]. Learning originates from experience, but its results do not affect single enclosed mental mechanisms (agent). Looking at the development of human beings as described in S. Freud's theory, adaptation and predestination affect many domains of the organism and bodily needs [Freud 1917, p. 346]. Cooperative learning in Multi-Robot systems is a relatively new field of research and as such challenged by complexity and dynamics [Panait 2005]. There is a high diversity of different approaches based on different assumptions. [Panait 2005] distinguishes the following types of learning in technical approaches:

- Machine learning
    - o Reinforcement learning
    - o Evolutionary computation
- Team Learning
    - o Homogenous learning
    - o Heterogeneous learning
    - o Hybrid learning
- Concurrent learning

There have been many approaches in the field of Reinforcement Learning, especially Q learning e.g. [Steward 2005], [Buffet 2001]. In general the learning in team allow either learning via communication, e.g. to share information, or without communication, by copying successful behavior of other team members. Although the model of this approach is designed to allow learning, this approach does not emphasize learning directly.

3.3.7 Robotic free will

A. Damasio describes two domains of free will in the context of Spinoza's theories in [Damasio 2003, p. 174]. Besides ethics, free will is coupled with the ability to make deliberate choices, and is thus a willful control of our behavior. Human beings are capable of evaluating actions and considering them as right or wrong [Damasio 2003, p. 174]. This goes along with the theory of M. Solms in [Solms 2002]:

Definition 3.33: The essence of „free will" appears to be the capability of inhibition – *the capability to choose* not *to do something [Solms 2002, p. 281].*

According to Solms the human mind possesses inhibitory mechanisms to suppress the stereotyped and evolutionary primitive behaviors encoded in the primary emotional systems [Solms 2002, p. 281]. This entails the question of free will in technical systems: Can a robot be conscious, feel and obtain free will (in a deterministic or semi-deterministic world)? J. Anderson deliberated his question in [Anderson 2002] and came to following conclusion. Yes, under certain preconditions it might be possible, and there is no heretical intention in doing so. However, M. Minsky challenges the existence of free will, describing it as a highly abstract structure depending solely on a set of fixed deterministic laws combined with a purely random set of accidents [Minsky 1988, p. 306]. This view appears quite pro-

vocative as dialectic is immingled with ethics and socially founded views. In our approach, a more abstract domain seems to be appropriate and I therefore follow the theory of J McCarthy in accordance with Damasio's description. For J. McCarthy [McCarthy 2005] the mechanism of "situation calculus" is part of the free will as it involves the consideration of consequences of alternative actions [McCarthy 2005]. In these terms, free will includes the capability of distinguishing between different actions, which is essential for intelligent behavior [McCarthy 2000a, p. 1]. Analogous to M. Minsky's suggestion that intelligence can be extracted from unintelligent material [Minsky 1988, p. 17], it is supposed that this assumption can be valid for free will as well. This entails that a set of rules and semantics with necessary constraints on decision making can fulfill the pre-conditions of emulating intelligent behavior (see for further detail Chapter 3.3.1). However, there is a major difference between being able to choose and being conscious of this ability. The former will not be discussed here.

3.4 Domains in building automation

According to [Kastner 2004], domestic applications are a major focus for future control networks in building automation. This goes along with the endeavors of the technical committee of the Building Automation, Control and Management of the Industrial Electronics Society (IEEE) [eBACM 2007], emphasizing approaches of inexpensive, open designs in the research of Field Area Networks. Building automation networks will subject to following requirements:

- − Vast number of control nodes

- − Robust physical channels

- − Physical dispersion of network structure

- − Flexibility in network management

- − Reduced cost

Besides these fundamental requirements on structure and technology of control networks, higher functionalities and services, e.g. security and safety, and scenario variations are expected [eBACM 2007]. Security: Buildings used for industry, offices or housing require additional applications of building automation like time and attendance, biometrics, access control, intrusion detection, necessary security arrangements to prevent intruders, abuse of functionalities, unauthorized access of information. As is the case with service robots in the domestic area, the cooperation with (existing) building and home automation systems is highly preferable for the benefit of both systems. These communication networks enhance the capabilities of service robots in team work, as they allow them to deliver information, which might not be perceived with the limited sensory facilities of the robot itself.

3.4.1 Distributed system

Definition 3.34: "...the term distributed system is referred to the set of autonomous processors and data storage devices, which are connected to a communication network"[Loy 2001, p. 43].

Definition 3.24 is tightly connected to the data transfer of shared data processing, as the information transfer of a shared communication network is essential for a distributed system in order to perform distributed tasks. According to Enslow, a distributed system can be described with five crucial characteristics [Loy 2001, p. 43].

- *Hardware architecture*
- *Principle of data processing*
- *Location of data storage*
- *Control mechanism on communication network*
- *System transparency*

These properties can also be used to define the level of system distribution (Figure 3.2). In this approach, system distribution is administered by a group of merely independent robots, building a multi-robot system (MRS) that has a loose coupling with the building system. Although the connection and communication with the building automation system has not been emphasized so far, some general system properties have been defined: The hardware architecture can be split in two parts: those used by embedded mobile agents, the service robots of the MRS, which are completely autonomous.
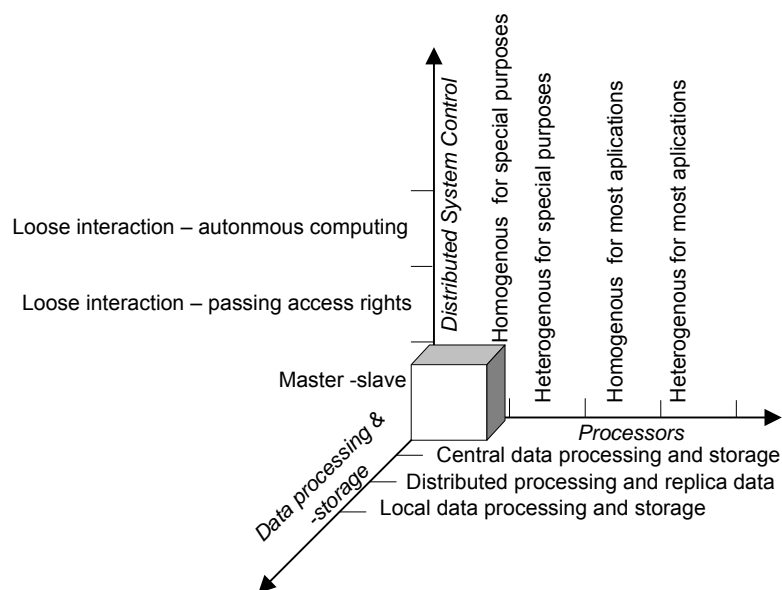
Figure 3.2: Valuation of distributed systems by Enslow (adapted from [Loy 2001, p. 44])

3.4.2   Control and monitoring

A technical process or machine requires monitoring functionalities besides safety controlling and security and optimization reasons [Loy 2001, p. 18]. While centralized systems posses controlling and monitoring functionalities in a unified software complex, they neglect the higher risk of logic errors due to faulty system design. Monitoring has to be proceeded through the whole operation time of handling process data and in case of a malfunction also output data. Designing monitoring functionalities without the knowledge of the control sequence increases the efficiency. In the case of building automation, the system interacts with a mainframe of control networks.

Figure 3.3 shows the basic concept of the overall system. It consists of a building automation system with an embedded multi-robot system (MRS), consisting of a team of (homogenous or heterogeneous) robots, which shall be coupled loosely with the building automation systems to share information and cooperate to achieve shared objectives. In this case, the MRS was defined as a group of two ore more robots, but it is also possible to use just a single robot as a solution for smaller sized projects and to reduce the overhead created by communication and coordination.
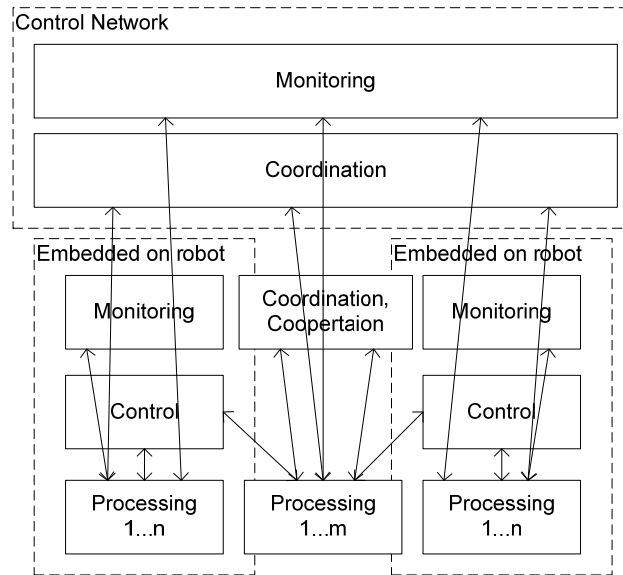
Figure 3.3: Hierarchical and parallel blocks of building automation system with service robots

In fact, the MRS and the control system are capable of operating independently, but share benefits, e.g. a higher amount of information (better overview of situations for the MRS, and therefore more adaptability), wider operation range (due to higher amount of direct interventions using robots as actuators of the building automatic system).

4 Approach

Definition 4.1: A model is „any structure that a person can use to simulate or anticipate the behaviour of something else" [Minsky 1988, p. 330].

At first glance, Artificial Intelligence (AI) and psychoanalysis seem to be two disciplines without a lot in common [Turkle 1989]. On closer inspection, AI proves to have a strong impact on psychology, as the way of looking at computers has affected the way of looking at the mind [Dreyfus 1989], [Papert 1989], and [Turkle 1989]. Although different in its structure and mechanism, there exists the hypothesis that the human mind and a computer system might share a common functional description on a high level of abstraction [Dreyfus 1989] and [Turkle 1989]. This goes along with the idea that a computer is "a device that generates intellectual behavior" [Dreyfus 1989]. However, AI is often implicated with a rationalistic perspective in which human experts are interpreted as data-processing systems possessing computer-like properties [Lueg 1997]. Psychoanalysis shows a radical view far more distant from rationalism, quantification and formalism [Turkle 1989]. But taking a closer look at the history and proposition of AI, AI has not emerged as a unified discipline but is characterized by a high diversity of different approaches [Papert 1989], [Turkle 1989], [Brooks 1991], and [Bryson 2007]. Although traditional approaches in AI tried to come up with a very objective and formalized representation of the world [Ratanaswasd 2005], [Lueg 1997], and [Turkle 1989], which has been accepted by and had influences on cognitive science and software engineering, and seems to be the predominant view until the end of the 20th century [Lueg 1997], another great number of approaches have since been made by new generations of scientists criticizing this methodology and emphasizing the flaws in traditional AI [Minsky 1988, p. 74], [Lueg 1997], and[Bryson 2007]. Inspired by bionics and psychoanalysis, new concepts and ideas on how to approximate humanlike behavior in technological systems were investigated in, e.g. [Pfeifer 1993], [Mochida 1995], [Lueg 1997], [Dodd 2005a], [Ratanaswasd 2005], and [Shell 2005], all testing the capability of diverse theories to overcome the limitations of classical AI. Nevertheless, the majority of these theories keep their focus on distinct human capabilities, aiming only for very specific tasks in a small range of applications [Bryson 2007]. In this frequently discussed method of designing AI, it has often been reiterated that a human, as a real person living and acting in a precise society, is far too complex for technological approaches [Singh 2003]. However, this shows a core problem of all works in AI. Before emulating the human mind, a comprehensive model of the essential and inevitable mechanisms of the human mind has to be composed. The artificial counterpart will still not reach the level of complexity of a human, covering the wide range of capabilities of a person, but it can be equipped with crucial functionalities that are accepted as unique and essential for the functioning of the human mind, thus designing an open extendable architecture. In order to achieve acceptable results in emulating mental mechanisms of the human mind, models of psychoanalysis, neuroscience and other disciplines of cognition science seem to be essential [Turkle 1989], [Sloman 2002], [Dodd 2005a]. Inspired by psychoanalysis, domains of natural and artificial intelligence go through changes towards new and more universal concepts and fundamental principles of the embodied mind. These promising approaches shall give a more unified picture of the main functionalities [Turkle 1989]. The goal of our approach is not to devise a new theory by valuing and extracting specific combined modules without considering the whole design.

The idea is to use universal models of cognitive science, which are based on consolidated findings on human behavior. Under the assumption that neuroscience and psychoanalysis are capable of describing and modeling the human mind in a sufficient way, their transfer into technical terms is highly desirable as the intellectual capabilities of the human brain represent an extremely efficient system, which cannot by far be achieved by current technological systems. Therefore, theories have been used that present the state of the art of human mind research in cognitive science in order to emulate key features of the human mind in a new behavior model. This entails the effect that the new technological model has to achieve either technological properness or properness in its assumptions, evaluated by another scientific area. Therefore, the goal of this thesis must not only be to improve automation systems, thus overcoming the flaws of current approaches, but also to demonstrate the correctness of the theories of another scientific discipline, which so far has been limited to the observation of individuals, under challenging, irreproducible circumstances.

Another very important aspect essential for this research is not only to investigate *how* the human mind works, but also *for what purpose* it works like that. What is the evolutionary goal of a human? Only in context with its goals can mental mechanisms be adapted and made use of for the changing goals in technological systems. Technical systems focus on completely different tasks as compared to living organisms. But the goals and requirements of technical systems do not necessarily contradict the goals in nature. E.g. survival strategies in nature are also an important issue for technology in order to enable systems to avoid critical states, which might lead to damages or failures. Besides the change of functional abstractions into technical analogous terms, another fundamental abstraction is essential. As all mechanisms of the human mind are embodied in an organism whose physiology is completely different from the conditions of technological systems, needs and qualities have to be redefined for technical systems, not unlike a number of mechanisms that have to be fine-tuned in individuals during their lifetime due to experience. The abstraction and transformation into technology has to be coherent, however, and must not violate basic rules of cognitive science.

4.1 Psychodynamics as a mastermind

Although the psychoanalytic perspective as proposed by Sigmund Freud [Freud 1923] and originator of young disciplines like psychoanalysis is meant to give a deep insight into the human mind, it has so far mainly been ignored in the discussion of intelligent behavior in traditional AI [Turkle 1989] and [Buller 2005]. Among others, the main reason for this neglect is the diversity of contradictory models and pictures apparently lacking even the least common denominator. This circumstance may cause disorientation in the validity and adaptability by giving as reason the "unscientific" image of these disciplines [Buller 2005] and [Turkle 1989]. This unfortunate prejudice has been enforced by the difficulty of methodically evaluating these theories. This may be a disincentive for the usage of psychoanalytic archetypes in emulating the human mind. Especially for engineers who are lacking years of experience in this area, even the variety of disciplines itself in cognitive science and their interdependence seems to be inscrutable.

However, due to recent developments, there has been great effort in the different disciplines for cooperation, aiming to give a more coherent picture that is based on mutual consent and can be adapted to the phenomena and evaluation processes of a greater group of disciplines. Based on a higher level of abstraction, concerning less the individual personality than a more unified, generalized picture of "the

human," it allows a better understanding and adaptability to the different disciplines of cognitive science, but also to technology. However, sexual properties will not be emulated in the functional model in this approach. The proposed research is founded on the basic assumptions of psychodynamics and its adaptation for machine psychodynamics [Buller 2005]. Psychodynamics must not be confused with psychoanalysis, which is founded on four general assumptions, namely [Buller 2005]:

− The evidence of unconscious processes taking influence on behavior

− The role of conflicts caused by opponent mental forces

− The existence of the Oedipus-complex

− The existence of sexual and aggression drives shaping and developing individual personalities

In our approach, only the first two assumptions, the core theory of psychodynamics, are considered for the functional model of this approach. As the physical robotic structure is not capable of reproduction, the mental capabilities of sexual affection are not required and will not be reproduced in the robotic behavior. Furthermore, the development of personalities coupled with individual experience shall be limited, as the robot cannot afford "lifetime" learning, starting on the very bottom to design its individual behavior model. Therefore, a minimum set of knowledge and a complete control architecture is highly desirable in order to reduce the introduction stage.

4.2 Methodology

Similar to Alan Turing's discussion about artificial intelligence [Turing 1950] more than 50 years ago, which is known under the name "Turing test", a new discussion, i.e. the question of artificial consciousness arises. One of the main criticisms of Artificial Intelligence is its failure to facilitate observable "consciousness", which is crucible for intelligence according to psychology and other disciplines of cognitive science. Besides the philosophical discussion that the consciousness, as well as intelligence, of a person, animal or machine cannot be "proved" other than by being this person, animal or machine, the principal argument against artificial consciousness, brought forth by diverse disciplines of cognitive science, is that consciousness is embodied and has been evolutionarily changed to serve the needs of the body, as organisms cannot sustain themselves [Solms 2002, p. 94]. This argument seems to be questionable, as a mobile robot is embodied as well, although this body possesses completely different properties from those of biological ones. Nevertheless, there are similarities in a biological and an electromechanical body:

− *Energy resources:* without energy, the robot is not capable of completing any missions or fulfilling tasks of any kind. The energy on a robot is limited. The lack of a minimum of energy, which is necessary to gain new energy to restore its functionality by itself, would cause "electrical death", although it does not necessarily lead to the physical destruction of the robotic body. A lack of energy has to be avoided at all costs. Considering cost functions evaluating the expected energy consumption of behavior arbitration is therefore an absolute must.

− *Damage:* as there are no mechanical repair mechanisms, damage always means stoppage. The risk of damage has to be minimized. The robot itself has to take care to remain intact while operating in missions so as not to limit its work capacity.

– *Machine fatigue and mechanical wear:* similar to damage and lack of energy, the slowly increasing mechanical wear and blackout of electronic parts due to long usage periods cause a decrease in utilization of the robot. Therefore, the breakdown of mechanical and electrical parts of the body has to be minimized. One preventive measure might be the monitoring of the load of the mechanical system and the capacity of the electrical system to ensure that there is no untimely loss through overcharge of the robot.

Similar to the organic body, it is crucible to preserve the robotic body, according to requirements of reliability, robustness, etc. Like fault tolerance, Artificial Consciousness (AC) is a method to cope with the high demands on future service robots. Therefore, psychoanalytical and neuroscientific theories have been reused in order to provide a better understanding of the mental functionalities, which are necessary for emulating higher functionalities of the human brain.

4.3 Obstacles and challenges

The main idea is to test the behavior architecture under simplified circumstances, creating thus a microcosm, which has neither a complex man-made environment nor a sophisticated social structure. The conjunction of such different disciplines may lead to "cultural" difficulties, e.g. terminology, methodic and research goals, which have to be reconsidered for this interdisciplinary work. It is difficult to find acceptance and cooperation in all scientific communities that are involved in an interdisciplinary work like the one endeavored in this approach. According to [Dietrich 2007], the following premises are inevitable to carry out a technical transformation of the psychoanalytically founded models:

– The functions of the human shall be convertible and modeled step by step with increasing complexity in technological design.

– Although subjectivity is crucial to the understanding of mental processes, it is only directly accessible to the subject itself. Therefore, subjectivity cannot be the objective of scientific research, which is restricted to a minimum of abstraction by making use of appropriate models as the basis of the approach.

– An interdisciplinary work like this requires mutual acknowledgement satisfying the requirements of both scientific fields, proving cautious examination of models and design.

– A consistent model is a necessity for the emulation or imitation of such a complex system like the human mind. It does not make sense to reduce or combine descriptions of distinct functionalities taken out of context.

– An interdisciplinary approach for transformation of models of the human archetype into technology allows only functional descriptions of higher abstraction. A technical system is not supposed to have the same physics as a human, using neurons, transmitters, etc., nor can it be expected that scientists in psychoanalysis or cognitive science have a deep understanding of computer science and control networks.

In analogy to the application of manpower to increase efficiency in a team, a similar concern arises in the assignment of robots when it is deliberated whether or not to form robotic teams for missions. One solution to build a technological behavior arbitration system in this context might be to construct and arm one robot with all physical and mental capabilities for all eventualities. But this leads to a clumsy

and expensive design, which needs too much resources and space for fulfilling the vast range of tasks. Instead of a solution where all functions are crowded into one single system, a number of more simply formed items cooperate successfully. There are several advantages by using robotic teams [Arkin 1998, p. 359] and [Sugawara 1999]:

−   *System performance:* as a task can be composed in steps exploiting parallelism in missions, the range of tasks and performance can be extended.

−   *Task enablement:* a vast range of tasks might require capabilities, which cannot be fulfilled by a single robot or only with a high level of specialization. In particular spatial separation and dispersion are a barrier for single-robot solutions.

−   *Distributed sensing and actions:* a robot team, possessing a higher number of sensors covering a wider field than a single robot, can explore unknown environment far more easily by sharing new information and adapting actions that might be highly distributed.

−   *Fault tolerance:* the redundancy of robots within a team can reduce the complexity of the robotic teammates themselves, which favor less error-prone solutions and increase the reliability of the system.

However, multi-robot systems are far more complex, which entails a range of challenges as follows [Arkin 1998, p. 360]:

−   *Interference:* the risk of blockages within the team by an increased number of robot-robot collisions seems to be inevitable and requires additional control methods to avoid mutual obstruction.

−   *Communication overhead:* communication is inevitable in a team and requires additional hardware, processing and energy, all of these resources that are quite limited on a robot and cannot be provided without cutting down other functionalities.

−   *Robustness:* communication by many devices in an unknown environment can suffer from noisy channels and electronic countermeasures. This uncertainty in communication challenges the coordination. Tasks allocation might become unclear due to poor communication.

−   *General costs:* even considering a more simply designed robot type for teams, the cost of two robots is usually higher than the one of a single robot.

These challenges are to overcome in a successful multi-robot design. The majority of multi robot systems are focusing on strictly homogenous robot teams, where each robot is equipped with the same set of functionalities [Parker 1998]. Future robot teams, however, shall contain specialized agents, which can fulfill a limited group of tasks. Thanks to the functional blocks and high abstraction of the software architecture, which allows for the configuration of basic actions, rules and images, various types of mobile robots can be equipped with the same basic functions of the control architecture, while the task mapping can be allocated under consideration of the individual specialization. As not only the physical structure of the robot can be split, also the behavior model itself can be varying. Showing the main impacts of functional blocks on the robot behavior, blocks can be separated and turned on/off.

4.4 Configuration and adaptation

To achieve progress in the design of a model of high complexity, the model has to be built in functional blocks, which can be removed, exchanged and configured, to estimate the importance and mechanism of its functions. In particular when using diverse control systems of different levels of complexity, the interplay has to be observed besides the overall, observable behavior selected for execution. As this model is founded on findings that come not only from the field of technology, but are also based on psychoanalytic and other interdisciplinary theories, a special apparatus to evaluate correctness and consistency of the theoretically designed functionality seems to be inevitable. Simulation allows a greater insight into the mode of operation than can be done by observation in the real world [Mondala 1993]. However, this goes along with parameterization, which is inevitable for operation:

– Optimized configuration of the model based on a chosen application

– Adaptation of the model due to the capabilities of the embodied agents

The model (described in Chapter 6.1 in detail) is described on a highly abstract level, containing all necessary mechanisms, but without pre-defined meta-data, e.g. all types of parameters, rules and restrictions which have to be adapted to the application and system properties[14], the model cannot be put to a test of its functionality. However, only the balanced combination of pre-defined inherited data and the functions of the model allow the successful design of efficient control architecture. A control architecture based on the model has to be designed with a minimum of parameters based on pre-defined assumptions. The purpose of simulation is not only to challenge the control architecture, but check its settings for optimization of the applied control architecture that inherits the model-based functions *and* the application-dependent configuration.

In order to achieve these objectives a high level of flexibility and scalability is required for the simulation environment and simulation tools. Especially the visualization of inner processes is a great challenge to adaptability, which will be discussed in the following chapters (Chapter 5.2 discusses current simulation programs and Chapter 7.2 presents the solution and design of this approach)

---

[14] The system can be a single robot solution, considering its abilities in physics, but also an overall control system in which the (embodied or computational) agent is embedded.

5 Preliminaries and related work

The proposed approach is based on different domains of computer science that focus on control architectures, robotic control and information theories and control networks, with focus on building automation and robotics, and inspired and influenced by concepts of psychoanalysis, especially psychodynamics. To emphasize the decisions and design goals of the proposed research work, a short introduction of major works preceding and influencing this approach is given below.

This approach is the result of close cooperation with other international researchers under the advice of experts in psychoanalysis. Although the influences from other researchers in this interdisciplinary work are characterized by a higher dispersion and diversity, the research work shall be grouped and summarized by their original domain for a better overview. The approaches in the following paragraphs will be categorized into the domains of building automation, robotics and other sciences. The aim is to give an overview of the scientific work of the various disciplines involved in this approach to present solutions that have influenced this approach, but also those which stand in contrast to this work. Therefore, different solutions related to the various aspects of this model shall be discussed below.

5.1 Trends in building automation

Communication control networks in building automation represent a major research area of the Institute of Computer Technology (ICT). In 1994, the Center of Excellence for Fieldbus Systems of Vienna University of Technology was founded together with the Department of Automation and the Institute of Electrical Measurements and Circuit Design. Under the direction of Prof. Dietrich, a large number of projects in cooperation with other research institutes and industries have been deployed, the results of which are presented in many scientific documents and books, e.g. [Dietrich 2006a], [Loy 2001], [Kabitzsch 2002], [Palensky 2003]. Prof. Dietrich is member of the technical committee for standardization, and he is Head of the Technical Committee of the Building Automation, Control and Management (TC BACM) [eBACM 2007] of the IEEE Industrial Electronics Society. Based on the profound specialist knowledge concentrated in the department, the developments in the research of an industry have been accompanied by delivering insight into the potentials and strengths of future developments and technology progress.

Control networks have become a very prosperous field in control and automation technology. Similarly, in robotics the first application field of control networks was the industrial automation [Kastner 2004]. This has changed in recent years, where control networks have gained increasing importance in home and building automation [Dietrich 2006a], [Kastner 2004]. Often condemned to a shadowy existence, this research field does provide tremendous benefits when compared to other networkings, e.g. Local Area Networks (LAN) [Loy 2001, p. xvii]. The automation relies on environmental information gathered from sensors and the capability to set actions through all types of activators [Loy 2001, p. xvii] and [Pratl 2006, p. 1], provided by sensor networks of increasing complexity, interconnecting different domains, e.g. heating systems, lightening and alarm systems connected by specialized proprietary systems, or the three main standardized players: LonWorks, EIB/KNX and

BACnet [Kastner 2004], which are mainly used in home and building automation. The main advantage of these open systems is that they present new technologies based on the knowledge of former field-busses. Providing the top level of technology without facing former problems in architecture and handling, these systems are considered to have a major influence on automation in the domestic area. Control networks considered to make allowance for barrier-free habitation founded the idea of an "intelligent environment" [Bruckner 2007, p. 1].

Looking at the developments of the last years, our society goes through changes in our way of life, communication, entertainment and interaction [Loy 2001, p. xvii], [Weiser 1991]. Similar to the developments in personal computers, which have started to become "invisible" and increasingly embedded in the environment [Weiser 1991], a lot of microcomputers take over tasks now without being noticed. This has a great impact not only on the technology, but also on the human as an operator [Weiser 1991]. Nevertheless, it shall be our objective to achieve less complexity in the "man-machine interface", as it cannot be expected that individual operators need special training to gain the necessary skills for interacting with these systems [Bolmsjö 2005]. Looking at the fast progress in control circuits, devices like microcomputers, nodes and sensors are becoming more efficient and economical at the same time [Loy 2001, p. 1]. Therefore, they are considered to be widely used, thus increasing the range of information dramatically [Pratl 2006, p. 1] and [Mahlkencht 2004,p. 1], widening the functionalities provided by the Internet. However, this must not affect the usability of the systems [Weiser 1991].

In order to meet these requirements, research in distributed networking and communication seems to be promising. In this approach, the communication and interaction with robots is of special interest in this context.

5.2 Robotic Control

Since the first robots have been introduced, robots of different sizes, shapes, raw material, sensing or locomotion systems have been designed. According to Definition 3.14 and Definition 3.17, a robot's main characteristic is its programmable, adaptable behavior. Therefore, the underlying control mechanisms to animate these robots are a major factor. As this research concentrates mainly on the behavior rather than on the physical capabilities of robots, mostly behavior-architectures shall be discussed here.

Furthermore, the preliminary work of this research, i.e. the designed robots, e.g. "Tinyphoon", a soccer playing robot, and intelligent solutions for building automation shall be discussed in the next chapters.

"The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions" [Minsky 1985, p. 163])

A great deal of robotic research concentrates on intelligent behavior arising from the interaction of the robot with its environment. There have been several methods used to develop intelligent behavior in robots. According to Arkin [Arkin 1998, p. 31], "the possibility of intelligent behavior is indicated by its manifestation in biological systems." Based on biological archetypes, various models of behavior-based robots have been built up. A classic approach related to robotic behavior is Artificial Intelli-

gence (AI). But using conventional AI based on the concepts of "sense-model-plan-act" has been argued to face several difficulties for mobile robots, e.g. the deficiency in dynamic environmental changing [Mochida 1995]. Most problems are found in extracting feasible information from a high complexity of sensory inputs and mapping the information to extensive inner models in order to be able to act. With a growing complexity of the internal models, the computational time is increasing dramatically. Though the computational ability of mobile robots has to be constricted, this instance leads to a bottleneck, which can lead to insufficiently slow behavior. This might be one of the reasons why the use of AI is limited to multi-robot systems where a coordinated behavior without external supervising and optimization is required. In the last years, there have been a lot of new approaches, e.g. behavior-based AI and emergent computation, providing robust behavior in dynamically changing environments. Robotics has troubles to achieve even simple animal capabilities. According to R. Arkin [Arkin 1998, p. 32], there are two major reasons for this dilemma:

− Biological systems use completely different hardware armed through evolution with functionalities that are not valuable for its silicon counterparts.

− A fundamental reason so far was that the knowledge about the functioning of the biological systems seemed to be inadequate and not adaptable to machinery.

However, information science is still finding new concepts in the disciplines of cognitive science. As biological information processing systems are able to provide feasible ideas for robotics, one of these data processing concepts, i.e. emotional decision making, has been introduced to enhance traditional behavior models. The human archetype demonstrates the flexibility and efficiency of decision making. One of the first concepts to implant "emotions" in their machines was implemented in a sensor input-driven steering system, proposed within the Braitenberg vehicles [Arkin 1998, p. 11]. This model is based on psychoanalytic theories of biological emotion and perception for the behavioral improvement of a new generation of robots. In the meantime, several approaches have been introduced using emotional concepts for decision making [Shibata 1996], [Mochida 1995]. But still, robots are not able to act in a "human-like" way of thinking [Singh 2003].

What is emotion? One of the main problems mentioned in the context of emotional behavior is that most scientists in AI lack a sufficient understanding of emotions [Pfeifer 1999]. Recently, excellent emotional models in neuropsychology [Panksepp 1998, p. 50] and [Damasio 1999, p. 67] and psycho-analysis [Freud 1923] and [Solms 2002] have been developed. This thesis delivers the theoretical basis for deploying a new behavior model to mobile robots.

Another important drawback in AI is that cognitive architectures and representations are often based on a single simple process, theory, or principle [Singh 2003]. Brooks' proposed architecture has the remarkable feature that it is based on several competence modules running in parallel, and deliberates which module is suitable. This principle can be found in the human architecture as well. The so-called control loops indicating (reflexive, reactive, routine, deliberative) behavioral pattern on different abstraction levels will also be a basic concept for this behavior model.

5.2.1 Cybernetics

Cybernetics is a composition of control theory, information science, and bionics. The objective of our approach is to use unified common control principles, which can be found in animals and have an

analogy to machines. The major idea is that the operation mode of an organism can be interpreted as that of a machine or otherwise a machine can work like an organism, using fundamental methods of mathematics for feedback control systems emulating the natural behavior. The focus of this approach is on the momentary situation developing a strong coupling of the machine (artificial organism) with its environment.

According to [Arkin 1998, p. 8], this research field has been founded amongst others with the approach of N. Wiener's theory in the late 1940s. N. Wiener's often discussed contribution ("Cybernetics or Control and Communication in "the Animal and the Machine," interpreted by R. Stichweh in [Stichweh 2004]) shows a clear concept of observable behavior that machines (defined as electromagnetic systems) can deploy to achieve similar results like their natural counterparts, but with different methodology. Emphasizing the cybernetic principles in robotic design, a further development has been made in the contribution of W.G. Walter, the "Machine Speculatrix", which has been embedded in the robot "tortoise" [Arkin 1998, p. 8] and [Holland 2001] and inspired following research, e.g. [Brooks 1985]. However, the arguably most famous example of cybernetics are the agents of V. Braitenberg, the so called "Braitenberg vehicles," which are based on a concept of inhibitory and excitatory influences and directly affect the motor systems of the robot. This inspiring work, introduced in 1984, has initiated a boom in recent research based on simple concepts, in the first place giving evidence that the complex behavior has its origin not necessarily in complex mechanisms [Mondada 1995]. A typical example of cybernetic behavior can be seen in Equations 5.1.

$$
\text{Stimuli} = \begin{bmatrix} (\text{movement} - \text{location}, 1.0) \\ (\text{enough} - \text{energy}, 0.7) \\ (\text{detected} - \text{object}, 0.3) \\ (\text{detected} - \text{enemy}, 0.0) \end{bmatrix} = \begin{bmatrix} s1 \\ s2 \\ s3 \\ s4 \end{bmatrix}
$$

$$
\text{Response(Stimulus)} = \begin{bmatrix} \text{move} - \text{to} - \text{next} - \text{object}(s1, s3) \\ \text{recharge}(s2) \\ \text{fetch} - \text{object}(s3) \\ \text{defend} - \text{object}(s3, s4) \\ \text{attack} - \text{enemy}(s2, s4) \end{bmatrix}
$$

$$
\text{Strength} - \text{of} - \text{Response} = \begin{bmatrix} 1.0 \\ 0.2 \\ 0.1 \\ 0.0 \\ 0.0 \end{bmatrix} \tag{5.1}
$$

The vectors, described in Equation 5.1, shows the stimuli driven stereo typed behavior, which can cover already a great variety of different fundamental behavioral pattern without sophisticated internal control mechanisms (the examples used in this equation goes along with an exemplary stereo type behavior described in Figure 7.4 used in Chapter 7. This might entail the question, why not to use so simple

5.2.2 Artificial Intelligence

"Computers and Thought *are the two categories that together define Artificial Intelligence as a discipline" [Brooks1991].*

One of the classic and so far most promising approaches in autonomous robotic control is Artificial Intelligence (AI). [Arkin 1998, p. 14] states the inception of AI with 1955 on the Darthmouth Summer Research Conference, where M. Minsky defined intelligent machines as a machine with a capability of creating an abstract model representing its environment in itself. As formal discipline is has existed for more than 50 years and in the beginning it was strongly influenced by computational architectures.

1985 saw the emergence of a new trend in AI, i.e. "autonomous agents research" [Maes 1994]. Based on this idea, robots should be able to find solutions to complex problems in robotic planning by using methods of knowledge and deliberative reasoning. AI concentrates on the relation of the representation of knowledge (e.g. world model) and the ability to plan future operational sequences. The characteristics of the human mind have been applied as algorithms in a computer-friendly way. A more or less flexible or efficient approach can be taken, depending on the established requirements, which have an influence on how artificial the intelligent behavior will appear.

Based on this knowledge, the machine shall be capable of finding and testing solutions based on its internal model before using them in the real world [Arkin 1998, p. 14].

5.2.3 Researches in action selection

Definition 5.1: "Action selection is the process by which any agent chooses at any instant, what to do next. [Bryson 2007]

Definition 5.2: A plan is "a sequence or partial ordering of behaviours which will attain the current goal" [Bryson 2000]

Starting as a part of AI, action selection (as defined in Definition 5.1) has become an important field of research in biology, as well as in other research communities [Bryson 2007]. With contributions from researchers such as [Minsky 1985], [Brook 1986 & 1991], and [Maes 1990], the computational models inspired by biological systems have emphasized the modular compositions towards behavior control (see Chapter 3.2.2) [Jones 2005]. However, according to [Bryson 2007] these models shifted the complex of problems from planning to integration, and are concerned with the question how to guarantee a coherent behavior in distributed parallel systems. There are different opinions in the community whether actions are discrete or not, and whether actions are application-independent or not [Bryson 2007], [Chapman 1985], and [Maes 1990]. Traditional approaches are based on a number of assumptions, which might not be generally valid [Bryson 2000]. For example, they neglect the fact that actions cannot be completely separated from perception as actions depend on expectations and context. Furthermore, many actions require constant feedback through perception [Bryson 2000]. [Bryson 2000] stated that the following design principles are essential for autonomous control architecture:

- Modularity allows simple, clear, and manageable control entities

- Hierarchical action selection simplifies conflict resolution and follows the principle of focus of attention

- Parallel monitoring systems allow the control architecture to be more responsive, capable of reevaluation and adaptation of priorities due to changes in attention

However, modularity might create other design difficulties as a straightforward hierarchy cannot be followed in functional decomposition. The risk of detecting dangers too late and leaving the agent unprepared is inherent in a centralized focus of attention [Bryson 2007].

In [Kortenkamp 1998], comprehensive software architectures for autonomous robots were introduced, giving a layered architecture and containing behavior modules and plans. However, examining the current state of the art shows that general control architectures providing complex (humanoid) behavior seem to be rare, and former leading researches seem to have no active development outcomes anymore [Bryson 2007]. However, with [diPaolo 2007] a two-layered architecture with behavior goal arbitration on the higher level has been introduced. Although not succeeding the concepts of layered models as proposed in [Kortenkamp 1998] and [Brooks 1991], this research gives some interesting action-selection mechanisms, selecting actions from a behavioral repertoire to emulate behavior of real orcas up to recent findings in biology and providing individual variations in behavior of specific animals in simulation as well. The model uses a finite state machine, which shall evoke need-driven behavior using memory and drives. This interesting approach aims to show behavioral semantics in different given situations. Using similar basic design principles like this approach, [diPaolo 2007] was not designed for robotic application in complex, changing environments. Therefore, environmental influences appear rather low and the current model seems to lack the necessary flexibility and adaptability that are required in real world applications.
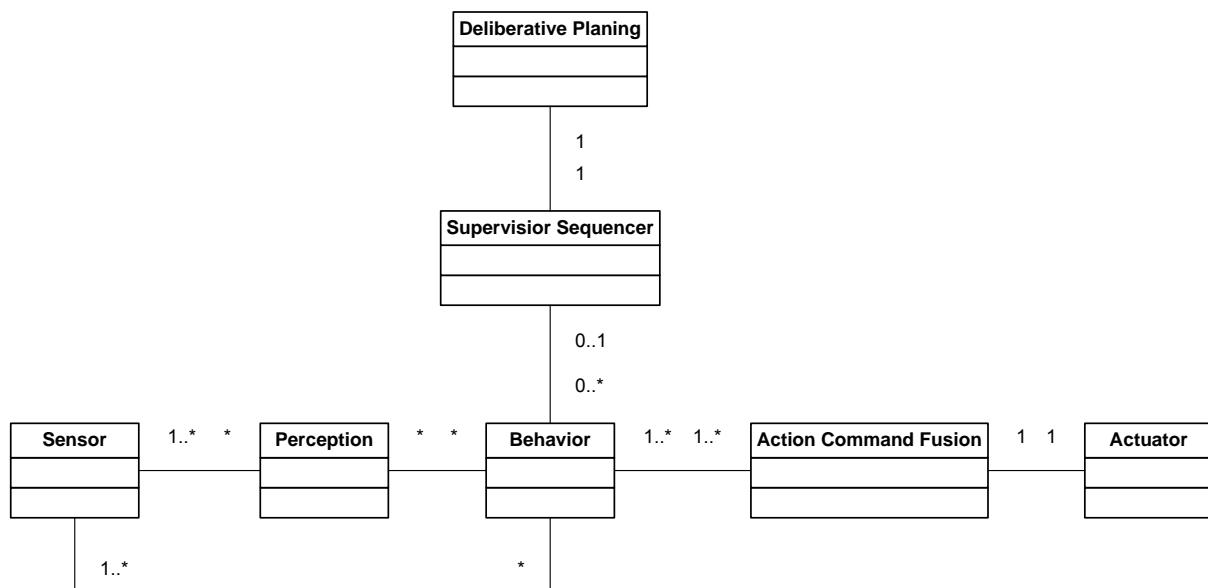
Figure 5.1: Generic hybrid architecture presented in the UML (adaptation of {Oerebäck 2003})

Beside a great number of application dependent and syntonized approaches a group of universal architectures providing unified platforms have been developed, and will briefly introduced in this chapter. In analogy to the survey of [Orebäck 2003], the basic requirements on a control architecture (not only for robots, but for any kind of system) can be outlined in terms of control, modularity, software engineering and run-time performance allowing conclusions about the potentialities of diverse general designs. A conclusion of the survey of [Orebäck 2003] is that most efficient architectures often rely on

a design based on a hybrid of reactive and deliberative control, using in general 3 layers, a deliberate layer, with focus on filling the gap of deliberative and reactive behavior, as shown in Figure 5.1.

Examples of architectures are, e.g. the AuRA architecture proposed by R. Arkin in [Arkin 1998, p. 215], which is based on a hierarchical structure, that inherits modules for mission planning, spatial reasoning, and plan sequencing. Other examples, e.g. 3T of [Bonasso 1997] or RAP [Firby 1989] use different number of layers ore do not directly match with layered structure.

A reason why overall behavioral concepts seem to have diminished in the last years might be that computer science, as has been proposed so far, may not be sufficient to facilitate intelligent behavior on its own. Depending on the methodologies, standard methods find it exceedingly harder to prove the correctness of a theory, and they are not well accepted in the AI community [Bryson 2007]. Sharing the hope of A. Bryson [Bryson 2007] that the research community might refocus, paying higher attention on methodology and general concepts, I hope this approach can be valuable contribution to this matter.

5.2.4 Multi-robot systems (MRS)

A general question in robotic design is whether a single-robot solution or a multi-robot system is more appropriate for specific applications. Various approaches, e.g. [Sugawara 1999], [Buffet 2001], [Parker 1998, 2002, and 2006], [Panait 2005] have concentrated on designs of multi-robot solutions and multi-agent systems respectively.

MRS systems are a major research field in Japan and provide interesting examples, among which the cellular robotic (CEBOT) system might have been one of the first [Arkin 1998, p. 361]. As stated by L. Parker [Parker 1998], the majority of MRS are merely focusing on homogenous robot teams, where each robot possesses an identical set of functionalities and movements. A reason for this approach might be advantages in reliability, as these all-purpose robots can be replaced more easily in case of failure, and their task allocation is less challenging. However, it is expected in the future that robot teams will contain different types of robots, which will be specialists in a certain group of tasks but therefore limited in their application. The Project ALLIANCE of L. Parker [Parker 1998] proposes a behavior architecture applicable for heterogeneous robot teams.

5.2.5 Simulation and experiments of robotic control

Investigations of single or collective behaviors of robots in the real world are time-consuming and very complex. Simulation environment can be a very efficient tool to validate an appropriate behavior of a robot or group of robots providing a simpler way of validation [Mondada 1994]. Although the programming environments and real-time visualization still had insufficient infrastructure ten years ago, there has been great progress in technology. Nowadays, numerous sophisticated, freely available and real-world control platforms provide excellent, inexpensive solutions for development [Blank 2004]. With the progress in the performance of contemporary hardware and software tools it seems appropriate that the statement of [Mondada 1994] that experimentation using standard infrastructures is hardly deployable, is outdated. The validation via simulation represents a potential way of validating control algorithms and architectures embedded in robots, promising success of these control systems in the real world. As emphasized in the survey of [Oberbäck 2003], it seems to be a general agreement that a commonly available software basis would be of great benefit, performing synthesis

of common architectures for technology transfer. However, the barrier of high diversity still remains, as the majority of robots posses solely proprietary application programming interfaces (API) [Blank 2004].
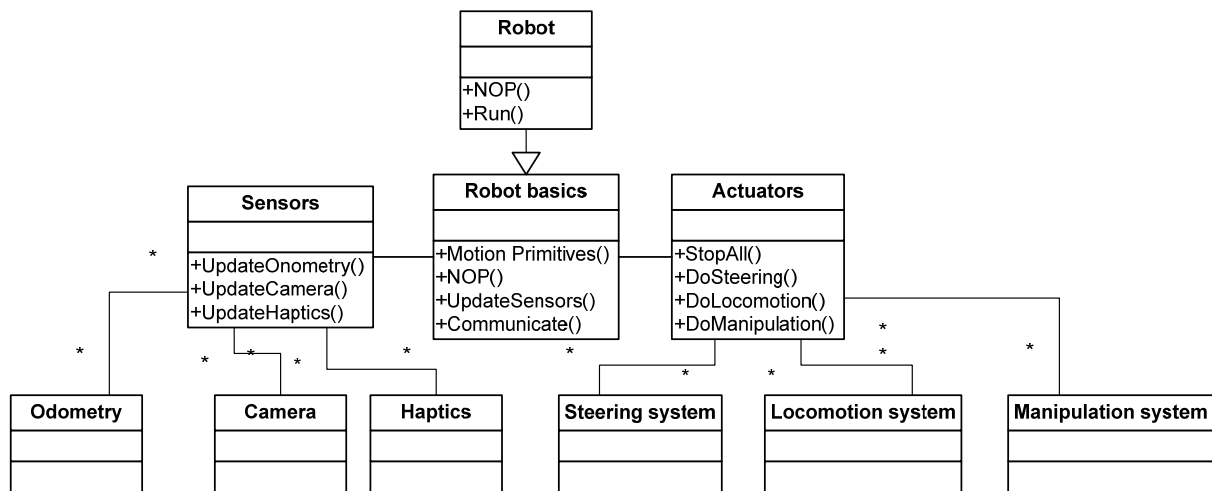

Figure 5.2: Example of a robotic structure

As emphasized in [Oberbäck 2003] theoretical models cannot be compared without implementations of architectures. Figure 5.2 shows an example structure for implementation as recommended in [Oberbäck 2003], containing a sample of sensors and actuators which have to be considered for implementation. Using universal simulation platforms allow a high variety of types and number of sensors or actors leading to different capabilities and motion primitives. While [Oberbäck 2003] considered archtitectures of Saphira, TeamBots, BERRA (Behavior based Robot Research Architecture), recent approaches, e.g. Mobile and Autonomous Robotics Integration Environment (MARIE) or Python Robotics (Pyro) etc. have been introduced. Table 5.1 gives a short overview of the main properties of these approaches:

Table 5.1: Overview of simulation tools (properties and evaluation capabilities)

	Saphira (6.2)	TeamBots (2.0)	BERRA (2.0)	Pyro	MARIE	AnyLogic (5.5)
Origin	SRI International Intelligence Center [Konolige 1996]	Based on ACE [Balch 2002] [Oerebäck 2004]	Royal Institute of Technology [Orebäck 2004]	Bryn Mawr Collage et al [Blank 2004], [ePyro 2007]	Based on ACE Université de Sherbrooke [Cote 2006], [eMarie 2007]	Xj Technologies [eAnylogic 2007]
Platforms	Linux MS Windows Solaris	Linux MS Windows Solaris MacOs	Linux Solaris MacOs	Linux MS Windows Solaris MacOs	Linux, MS Windows	Linux MS Windows
Program Language	C	Java	C++	Python C/C++	C++, XML	Java, XML

				interface		
Free License	No	Yes	No	Yes	Yes	No
GUI	Yes	No	Yes	No	Yes	Yes
Sensor extension	No	Good (JNI)	Very good	Yes	Yes	Good (JNI)
Hardware abstraction	Good	Very good	Very good	Good	Good	Very good
Documentation	Manual	Web, class diagram, web, book	Class diagrams, web	Web, publication	Web, publication	Web, manual

In general simulation tools are providing several forms of sensors and robotic properties, allowing a high variety of robots to be used. Some approaches, like Pyro are designed also to teach students and therefore allow a very abstract and functional oriented description.

Another interesting approach was emphasized by the a group at the Université de Sherbrooke, Canada [Cote 2006], and [eMarie 2007]: to cope with the diversity of interoperable software used in robotics a middleware framework to integrate commonly used software and frameworks like CARMEN, Player, RobotFlow, etc. is provided.

In contrast to the simulation, F. Mondada's approach [Mondada 1994, 1995] was to design a miniaturized robot Khepera of modular design that facilitates the change of sensors and actuators. Using a versatile software structure allows debugging of algorithms and visualization of experiment results [Mondada 1994]. However, this is still a proprietary solution that does not provide any adaptability.

5.3   Cognitive disciplines and psychoanalysis
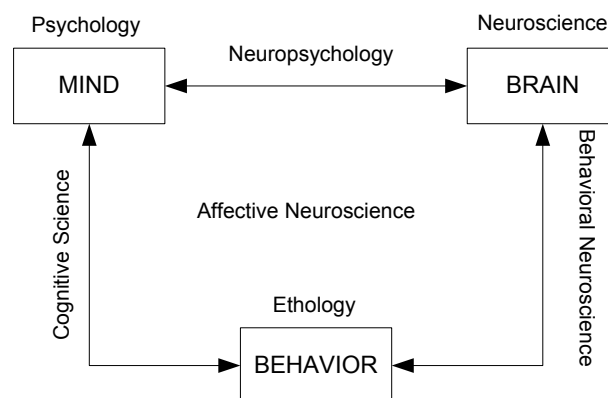
Figure 5.3: Relation between disciplines of cognitive science (based on [Panksepp 1998, p. 31])

In no aspect of the human mental life is it more important to understand the quality and meaning of the genuine human archetype than in the case of emotions [eStanford2007]. Many scientists have tended to neglect them, as the word "emotion" seems to cover a variety of phenomena. Facing a reverse trend in the last decades, emotions became the focus of strong interest within cognitive science. In psycho-

analysis systematically introduced and described in their functions, emotions have also gained major interest in other disciplines, particularly psychology, but also neurology and evolutionary biology.

J. Panksepp introduced in [Panksepp 1998] the relation of different disciplines in cognitive sciences (Figure 5.3). The goal of cognitive scientists is to identify the key mental representations and to try to figure out how they work. The following paragraph describes the assumption of major emotion theories and their consequences, as well.

According to psychoanalysis, the highest goal of all animate creatures is the "survival to reproduce" [Solms 2002, p. 19, 29]. This means that on the most fundamental level the essence of all human behavior is based on the motivation, a virtual impulse to provide the organism with all necessities to sustain its vital functions. The principal task of the brain, in similarity to any control system of an embodied autonomous agent, is to mediate the divide of two 'worlds': the internal world, the internal milieu of the body, and the external world, the external environment to focused on increase the potential of survival. The adaptation of this principle is a core topic in Chapter 6.2.1, where its impact on this research will be discussed in detail. The internal world of the body has to interact with the external world that surrounds it in order to ensure that the world meets its needs. This task is managed by the brain [Solms 2002]. The mental apparatus, in the case of humans the human brain, is linked with the environment via a sensory apparatus and a motor apparatus, which enable it to receive information about the environment from the environment and to act in the world that surrounds the body. Considering initialization, neither the brain nor the mind is completely genetically determined or fully pristine [Damasio 1994, p. 111].

5.3.1   Emotion theories

The human brain and its mental processes are one of the most complex systems that can be found in biology. Due to the high demands on fault tolerance**,** availability and reliability, organisms need far more efficient mechanisms to be successful in the process of evolution. Therefore, brilliant apparatuses like the human brain have developed and shall be used as an example for the technical models of mobile robots, which face similar problems within the application field of building automation.

Evolutionary, complex cognitive control can be found in higher animals, which are not simply passive reflex machines containing stereotyped responses based on environmental stimuli [Panksepp 1998, p. 38]. In animals, the evolutionary ancestors of the human being, the first adaptive behavior emerged, which is generated spontaneously and flexibly initiated through recognition of internal and environmental events. The flexible neural circuits of animals can be interpreted as "master routines," covering various sub routines with the necessary intrinsic behavioral flexibility, which has brought to the human concept the emotional command system connecting events and their meaning.

Emotions are fundamental in human behavior. With the assistance of emotions, the selection and rating of necessary actions within situations of high complexity and unreliability can be designed. Being equipped with emotional behavior allows a distinction between "desirable" and "harmful" action patterns for a variety of situations without the necessity of having experienced them before.

No mental processes without emotions

R. Plutchik stated in [Plutchik 2001], that emotion is a complex chain of coupled events of different hierarchy. Beginning with a stimulus, emotions initiate psychological changes, impulses to action, and

specific goal-directed behavior. The human behavior and human mental structure are bound on cyclic mechanisms of emotions. These mechanisms themselves are not necessarily linked with the human consciousness, but consciousness strengthens the impact of feelings [Damasio 1999, p. 57].

− Consciousness is a mental apparatus containing two main issues: the connection between organism and object and the change of the organism due to the presence of the object [Damasio 1999, p. 33].

− Feelings are the private mental experience of an emotional state. They are perceptions of having an emotion.

− Emotion is the public reaction of a feeling.

Similar to drives, emotions have their evolutionary origin in the regulative processes of homeostasis. Emotions are a complex bundle of chemical/ neuronal reactions building a pattern for regulative features. These concepts can also be found in animals. Emotions are used automatically (unconsciously) to produce fast reactions in urgent situations. They are a mini-concept of a set of reactions initiated by a range of stimuli. The structures building theses emotions are locally limited to some specialized structures, the emotional systems. The human brain has a representation of these sets of emotions on the conscious surface in the form of so-called "feelings".

All images (defined in *Definition 3.3*) based on external stimuli or recalled by memory are emotionally labeled and valued. Through living in complex social structures, the process of conditioning allows an individual to link neutral objects and situations with emotions, which have originally been used for other objects and situations (which are naturally linked with emotions) [Damasio 1999, p. 75]. This is one of the most important and effective forms of learning.
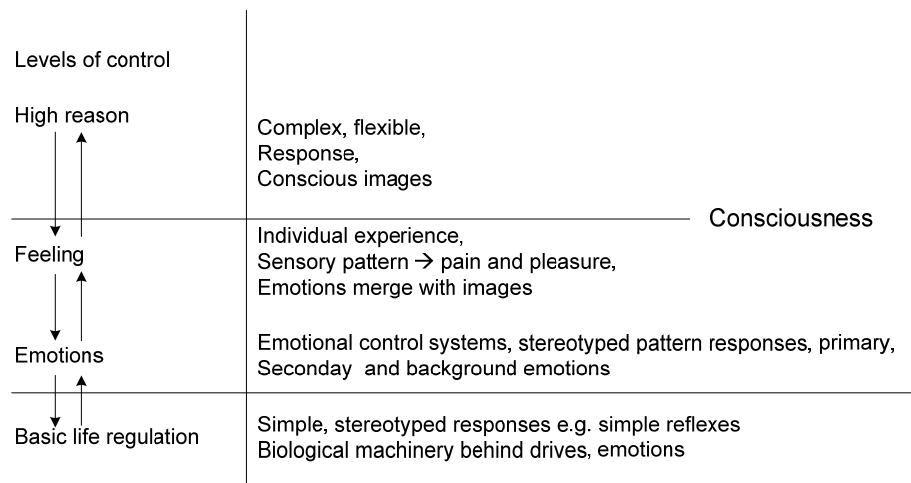

Levels of control

High reason → Complex, flexible, Response, Conscious images

Consciousness

Feeling → Individual experience, Sensory pattern → pain and pleasure, Emotions merge with images

Emotions → Emotional control systems, stereotyped pattern responses, primary, Seconday and background emotions

Basic life regulation → Simple, stereotyped responses e.g. simple reflexes Biological machinery behind drives, emotions

Figure 5.4: Levels of control in a human (adapted from [Damasio 1999, p. 55])

As shown in Figure 5.4, theories like those of A. Damasio emphasize major relations of feelings, consciousness and emotions, which have been discussed in former chapters (Chapter 3.1.5 gives a general introduction into how emotions are defined and used in this approach, while Chapter 3.1.6 shows the relation to higher mental functions, which are partly considered in this approach)

In case of emotions, there are various theories about the classification of emotions (an overview is shown in The term behavior in psychoanalysis means a sequence of processing steps, as emphasized by R. Plutchik and illustrated in Figure 5.5. According to Plutchik's theory [Meyer 1999],

[Plutchik 2001], every individual is armed with groups of basic actions, which can be directly linked to emotional systems. There are numerous approaches that conform to this view (some of them are presented in some detail in Chapter 5.2.3 on computer science, using a similar way of data processing for their control system). The theory that the output (directive to choose an action) is determined by the input (perception and cognition of environmental data) and the system state comes close to a computational theory of finite state machines (FSM), which uses a mathematical concept to describe behavior as a composition of a finite number of states, transitions and actions.

). The majority of psychoanalytic and cognitive theories distinguish between two main types of emotions: primary and secondary emotions (see Chapter 3.1.5 for further detail). A very prominent version is Antonio Damasio's classification, which resembles other theories. He defines three different types of emotions in [Damasio 1999], covering the whole range of emotional states of the human mind:

- Universal, primary emotions: According to Antonio Damasio, there are 6 primary (basic) emotions in the human [Damasio 1999, p. 67] which are the building blocks of secondary emotions:

    o Happiness

    o Pleasure

    o Grief

    o Anger

    o Surprise

    o Disgust

- Social, secondary emotions: These emotions are built up by social structures. Although they will not directly be used in this model, they are indirectly presented by the rules of the Superego.

- Background emotions: unlike the other two types, background emotions define long-term emotional states, stressing and distressing actual (situation and external stimuli-dependent) emotions.

Beside a comprehensive set of primary emotions, the extraordinary part of this theory is that A. Damasio adds a third category of emotions, i.e. background emotions, which appear as a major factor in the time-dependent process. A mental process leading to a decision might not depend solely on the given situation, but also on its previous situation.

The term behavior in psychoanalysis means a sequence of processing steps, as emphasized by R. Plutchik and illustrated in Figure 5.5. According to Plutchik's theory [Meyer 1999], [Plutchik 2001], every individual is armed with groups of basic actions, which can be directly linked to emotional systems. There are numerous approaches that conform to this view (some of them are presented in some detail in Chapter 5.2.3 on computer science, using a similar way of data processing for their control system). The theory that the output (directive to choose an action) is determined by the input (perception and cognition of environmental data) and the system state comes close to a computational theory of finite state machines (FSM), which uses a mathematical concept to describe behavior as a composition of a finite number of states, transitions and actions.

gives an overview of the different terms used in different emotion theories in psychoanalysis, neuro-psychology and cognitive sciences. The idea of primary emotions as described in earlier chapters (see in particular Chapter 3.1.5) is quite common, but the name and number of emotions vary from one

theory to the next [Lorenz 2007a]. According to A. Damasio, there are 6 primary (basic) emotions [Damasio 1999, p. 67], which is a similar concept to the classification of G. McDougall (summarized in [Meyer 1999]). But both theories do not consider the degree of the pressure these emotions might contain, which R. Plutchik has emphasized as a three-dimensional circumplex model[15], which is illustrated in its basic outline in The term behavior in psychoanalysis means a sequence of processing steps, as emphasized by R. Plutchik and illustrated in Figure 5.5. According to Plutchik's theory [Meyer 1999], [Plutchik 2001], every individual is armed with groups of basic actions, which can be directly linked to emotional systems. There are numerous approaches that conform to this view (some of them are presented in some detail in Chapter 5.2.3 on computer science, using a similar way of data processing for their control system). The theory that the output (directive to choose an action) is determined by the input (perception and cognition of environmental data) and the system state comes close to a computational theory of finite state machines (FSM), which uses a mathematical concept to describe behavior as a composition of a finite number of states, transitions and actions.

. This model gives a very interesting schema of the overlapping and interlacing of emotions, which can also be used as guideline on the introduction of complex emotions. Comparing the proposed different classifications of emotion, the least common denominator seems to be fear and anger, which are easily observable, and the majority of theories propose a form of grief or sadness, respectively, which J. Panksepp subsumes as a panic system that invokes separation distress. An interesting result is that in terms of positive emotion it seems there is no agreement in their classification (The term behavior in psychoanalysis means a sequence of processing steps, as emphasized by R. Plutchik and illustrated in Figure 5.5. According to Plutchik's theory [Meyer 1999], [Plutchik 2001], every individual is armed with groups of basic actions, which can be directly linked to emotional systems. There are numerous approaches that conform to this view (some of them are presented in some detail in Chapter 5.2.3 on computer science, using a similar way of data processing for their control system). The theory that the output (directive to choose an action) is determined by the input (perception and cognition of environmental data) and the system state comes close to a computational theory of finite state machines (FSM), which uses a mathematical concept to describe behavior as a composition of a finite number of states, transitions and actions.

). Based on neurological findings, J. Panksepp's theory of the seeking system, which is directly linked with the internal "need states" [Panksepp 1998, p. 168] in analogy to drives (Definition 3.8) in S. Freud's terminology, gives evidence and shall dominate the approach of this thesis.

5.3.2   Behavior determination in psychoanalysis

The term behavior in psychoanalysis means a sequence of processing steps, as emphasized by R. Plutchik and illustrated in Figure 5.5. According to Plutchik's theory [Meyer 1999], [Plutchik 2001], every individual is armed with groups of basic actions, which can be directly linked to emotional systems. There are numerous approaches that conform to this view (some of them are presented in some detail in Chapter 5.2.3 on computer science, using a similar way of data processing for their control system). The theory that the output (directive to choose an action) is determined by the input (perception and cognition of environmental data) and the system state comes close to a computational

---

[15] According to [Meyer 1999] this model was introduced in 1980 for the first time. A description can be also found in [Plutchik 2001].

theory of finite state machines (FSM)[16], which uses a mathematical concept to describe behavior as a composition of a finite number of states, transitions and actions.

Table 5.2: Overview of primary emotions introduced in different emotional theories

Antonio Damasio	Robert Plutchik			Graham McDougall	Jaak Panksepp (Mark Solms)	
Happiness	Ecstasy	Joy	Serenity			
	Admiration	Trust	Acceptance			
Fear	Terror	Fear	Apprehension	Fear	Fear system →	Anxiety, Foreboding, Alarm
Surprise	Amazement	Surprise	Distraction	Surprise		
Sadness	Grief	Sadness	Pensiveness		Panic system→	Grief, Loneliness, Seperation Distress
Disgust	Loathing	Disgust	Boredom	Disgust		
Anger	Rage	Anger	Annoyance	Anger	Rage system→	Hate, Anger, Indication
	Vigiliance	Anticipation	Interest		Seek system→	Anticipation, Hope, Desire
				Elation		
				Subjection		
				Affection		

A major question seems to be whether it is legitimate to reduce the complexity and variety of behavior for a technical model, using a limited amount of discrete states and an endless number of actions. R. Plutchik defined the following fundamental adaptive action categories:
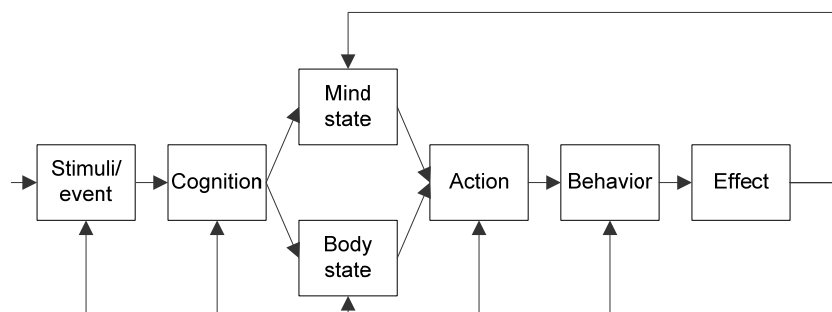

Figure 5.5: Processing in the mind

· *Protect (part of fear control system):* to prevent the robotic body from external and internal dam-age. Behaviors like retreat or escape shall increase the distance between the robot and the source of danger.

· *Destroy* (*part of rage control system*): to abolish obstacles, which prevent the robot or organism to fulfill an important need.

· *Reattach (part of panic control system)*: this type of feeding shall rebuild the physical power of the robot.

---

[16] A general introduction to this theory can be found e.g. on http://en.wikipedia.org/wiki/Finite_state_machine or other Internet pages.

- *Mutual support (part of panic control system)*: this shall improve cooperation to restore the ability to reach shared goals, e.g. defense and improve survival.

- *Reject (part of fear control system)*: this shall help to avoid damage to the robotic body.

- *Examine (part of seek control system)*: this shall help to increase the information about the environment.

- *Orientate (part of seek control system)*: to achieve a broader knowledge about the environment, the agents have to wander around and observe.

Table 5.3: Classification based on Plutchik's proposed innate behavior categorization

Stimulus/ event	Cognition	Feeling state	Overt behavior	Effect
Threat	Danger	Fear	Escape	Safety
Obstacle	Enemy	Anger	Attack	Destroy obstacle
Gain of value object	Possession	Joy	Retrain and repeat	Gain resources
Loss of value object	Abandonment	Sadness	Cry	Reattach to lost object
member of one's group	Friend	Acceptance	Groom	mutual support
Unpalatable object	Poison	Disgust	Vomit	Eject poison
New territory	Examine	Expectation	Map	knowledge of territory
Unexpected event	What is it?	Surprise	Stop	Gain time to orient

Similar to the concept proposed in [Kitamura 2002], behavior can be situated on different hierarchy levels: Complex actions can be split into basic actions, which in themselves are not emotionally labeled. Only their combination with higher order actions, which are goal-directed, can be emotionally labeled and grouped for pre-selection in decision making. The complex actions can be grouped into four emotional categories as proposed in [Panksepp 1998, p. 53].

5.3.3   Freudian psychoanalysis

Numerous approaches in different disciplines try to describe the human mind and its functionalities [Beers1996]. Due to the different applications, e.g. healing of a concrete mental disorder of an individual patient, diverse theories have emerged focusing on different areas. Therefore, a selection of theories that seem relevant to this research shall be introduced in the following section.

Ambivalences of the mind

S. Freud introduced in [Freud 1917] a series of theories with the intention to "clarify and carry deeper the theoretical assumptions on which a psycho-analytic system could be founded" [Freud 1917 p. 222]. Emphasizing the active, passive, and reflective expressions of drives, which give rise to the "subjectivity", i.e. ego, Freud introduced three main polarities of psychic life [Freud 1915]:

- Real polarity: Ego – object (external reality): this polarity allows the differentiation of internal and external reality and is based on the concept of drives, which are inescapable, while the external reality, which comprises subtle stimuli, can be avoided.

- Economic polarity: pleasure – unpleasure: the opposition of this pair allows an explanation of hate and love, as it indicates the continuing pressure of a drive and the relief of it when it is fulfilled.

- Biological polarity: passivity – activity: Based on the idea of subject and objects, this polarity expresses the ambivalence of drive impulses.
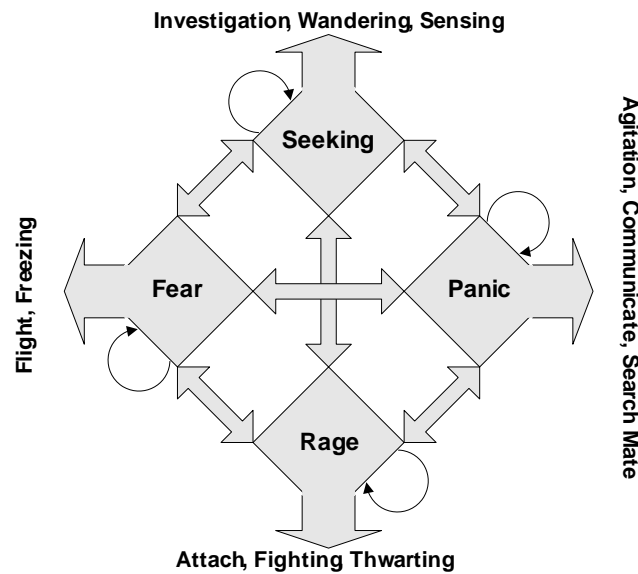

Figure 5.6: Behavior arbitration adapted from the biological archetype of [Panksepp 1998, p. 53]

Using these polarities, Freud tried to analyze phenomena like sadism/masochism, voyeurism/exhibitionism, and love/hate as pairs of opposites. Although the application itself is outside the focus of this research, the basic assumption that oppositions have their origin in polarities, which are neither predefined, nor predictable, appears valid for this research. The theory also implies that behavior is not entirely predictable as there is no evidence that similar external (and even internal) influences may lead to the same pre-defined behavior. However, in Chapter 6.2.4, I will emphasize that there is a certain possibility to categorize behavior, and assume action tendency according to a given situation.

Psycho Dynamic Theory

A second important contribution of S. Freud, the psychodynamic theory, is based on the idea that a large part of the human mind is unconscious and that the contents of the unconscious (the unthinkable) are the source of a great deal of our motivation. As the word dynamic implies, it suggests further concepts of Freud: drive force or motivation (Chapter 3.1.4). In [Freud 1923] Freud introduced a theory about the forces of the psyche, using theories of thermodynamics as a metaphor to explain human personality [Freud 1923]. Using a tripartite division of the human mind, he introduced the Id, the Ego, and the Superego:

- The Id is described as summary of biological needs and drives. Its function is to provide energy for the system, analogous to fire, which provides energy in thermodynamics.
- The Superego is seen as a set of society's rules, the so called "voice of conscience". Transferred to thermodynamic terms, the superego would be situated on the top of the apparatus that transforms the water it contains into steam.
- The Ego represents the conscious mind that contains thoughts, judgments and memories. According to the thermodynamic metaphor, the ego comprises the parts of the apparatus, e.g. the wheels released by the steam.

Using this analogy allows Freud to describe the idea of personality, which is not observable with a tangible model of energy heating water and steam. Using the first law of thermodynamics and apply-

ing it to the psycho-dynamics theory, it states that: "Energy can neither be created nor destroyed". Adopting this to the human personality, Freud hypothesized that "psychic energy" can be neither created nor destroyed. It just can be converted from one form into another form.

5.3.4 Social psychology

Social psychology and behavioral sciences lack the experimental methodologies found in natural science regarding basic functionalities of the human organism. Considering lower animals, only very specific, primitive functions can be observed, which do not necessarily add up to human basic functions. To coordinate the observed primitive functions into a meaningful concept and understand the basic characteristics, the main abilities of the human mind have to be separated and classified. Reexamining the arbitrary divisions within psychology, it becomes apparent that the core functionalities like learning, motivation, emotion, etc. are inevitable but hard to put to test as separate values.

Therefore, experimental psychology faces the dilemma of evaluating humanlike behavior. This means additional challenge for science in its efforts to evaluate behavior based on architectures emulating the human mind. The major question is in which experimental situation can the unique and efficient capabilities of the human be extracted and tested? How can we make them visible, comparable and ratable?

Criticizing experimental psychology, the Japanese psychologist M. Toda showed an interesting approach about the observation of mental functions which is summarized in [Toda 1982]. To deliberate the basic and efficient methods of the human mind for problem solving, he proposed game-like situations creating a microcosm of problems invoking mutually dependent problems, where the subject, man or machine, cannot survive without using major basic problem-solving strategies. As in a game-like situation, where problems have to be solved and opponents show up that have to be dealt with, is considered to be a sufficient test environment to uncover basic characteristics of human interaction, this test bench meets the requirements for a systematic qualitative evaluation of control architectures emulating the same basic characteristic behavior of human. In analogy to the Turing test, a human playing this game should find the same or similar solutions as the robot equipped with a control architecture functional concepts like the mind. As Toda tried to encourage humans to pretend to be more simplified robotic subjects, robots here shall try to act as humanoid as possible. In this context, a type "Turing test" for artificial consciousness can be provided.

Within the research of T. Mochida et. al [Mochida 1995], a control architecture for autonomous agents based on the Braitenberg architecture has been designed. The robot can have two states expressing pleasantness and unpleasantness, which is rated in a state variable called "frustration". Emotions discriminate and identify the current relationship between the autonomous agent and its environment. This emotional method is used merely for goal pursuing and shall help the robot to escape traps.

An approach dealing with emotions in robotic behavior was carried out by R. Pfeifer [Pfeifer 1994], [Pfeifer 1996], and [Pfeifer1999], implementing the first steps of an emotional behavior model based on the approach of the Japanese psychologist M. Toda [Toda 1982]. Toda developed the scenario of the "Fungus Eater" not only to examine emotional, but also intelligent behavior in general. Instead of using traditional methods of cognitive psychology, Toda's model contains abstract instructions to design an autonomous system, the "fungus eater" robot which reproduces humanoid behavior in a simplified, artificial real-world environment. Another emotionally driven behavior model was proposed by T. Shibata et al. [Shibita 1996]. It is based on a model of the neuro-psychologist and psycho-

analyst M. Solms. The model also uses frustration for rating the behavior of team mates. This mechanism shall provide a higher level of cooperation within a team of robots.

The approaches presented above deal with the key factors of emotional intelligence. However, there is a significant lack of a complete description of emotional decision making and how it relates to rational decision making. However in both theories, the Turing test [Turing 1950] and the social Fungus Eater of [Toda 1982], intelligence is supposed to be determined by the observable appearance of intelligent behavior. The major question arises is, is this a valid approach and which assumptiosn are necessary for the validity in this context. This will be discussed in Chapter 8.

5.4 Preliminary work - a brief history

Based on the discussions and approaches presented above on the Institute of Computer Technology, Vienna University of Technology, numerous projects have been initiated and deployed in the last years, providing the scientific background on which this thesis is built. The following chapter gives a brief introduction to the main projects, showing the history of this model and its future purpose and succeeding work.

Originally, design idea is to evaluate modern concepts in building automation meeting and considering future trends by introducing new design concepts founded in nature [Dietrich 2000]. Focusing on the objectives of reliability, fault-tolerance and performance in control, a principle design goal was an open control network that allows the concentration of vast information of different industries and domains. Using bionic concepts [Brainin 2004], the benefits for other applications have become apparent. This idea has been introduced by D. Dietrich in [Dietrich 2000], initiating diverse approaches derived on the Institute of Computer Technology. Numerous ideas and concepts have been introduced in the last 8-10 years, covering further fields of application and affecting more technological domains, e.g. robotic control [Roesener 2006], [Novak 2006] [Deutsch 2006].

At first, this research was inspired by domains of neurology and neuropsychology, and especially by the contributions of O. Sacks [Sacks 1986] and V.S. Ramachandran [Ramachandran 2004], which gave fascinating insight into fault tolerant concepts of the human body that provided a range of interesting concepts for networking and information processing. The first model, the so called "Perceptive Awareness Model," started in 1999, emerging from the SmartKitchen (SmaKi) project, which initiated further approaches (Figure 5.7). An introduction will be given in Chapter 5.4.1.

In 2003, a new research group was founded by Prof. Dietrich, called "Artificial Recognition System" (ARS), which succeeded the first project, SmartKitchen (SmaKi). The ARS project intended to enhance automation control with methods of human perception and situation recognition for building automation [Dietrich 2004a, Dietrich 2004b]. Inspired by the hypothesis of the Russian neuropsychologist A. Lurija, the first endeavor of the project was to improve the perceptive capabilities of automation systems, thus creating an abstract representation of the world [Pratl 2006, p. 15]. The concept of symbolization shall improve the performance in control, based on redundant and distributed information. The enhanced design for building automation allows the evaluation and classification of situations and facilitates the supervision of human activity, which can be beneficial for services like object or person tracking, surveillance, comfort and rehabilitation. However, the active support, which is particularly crucial in the last field of application, is not considered in this approach.
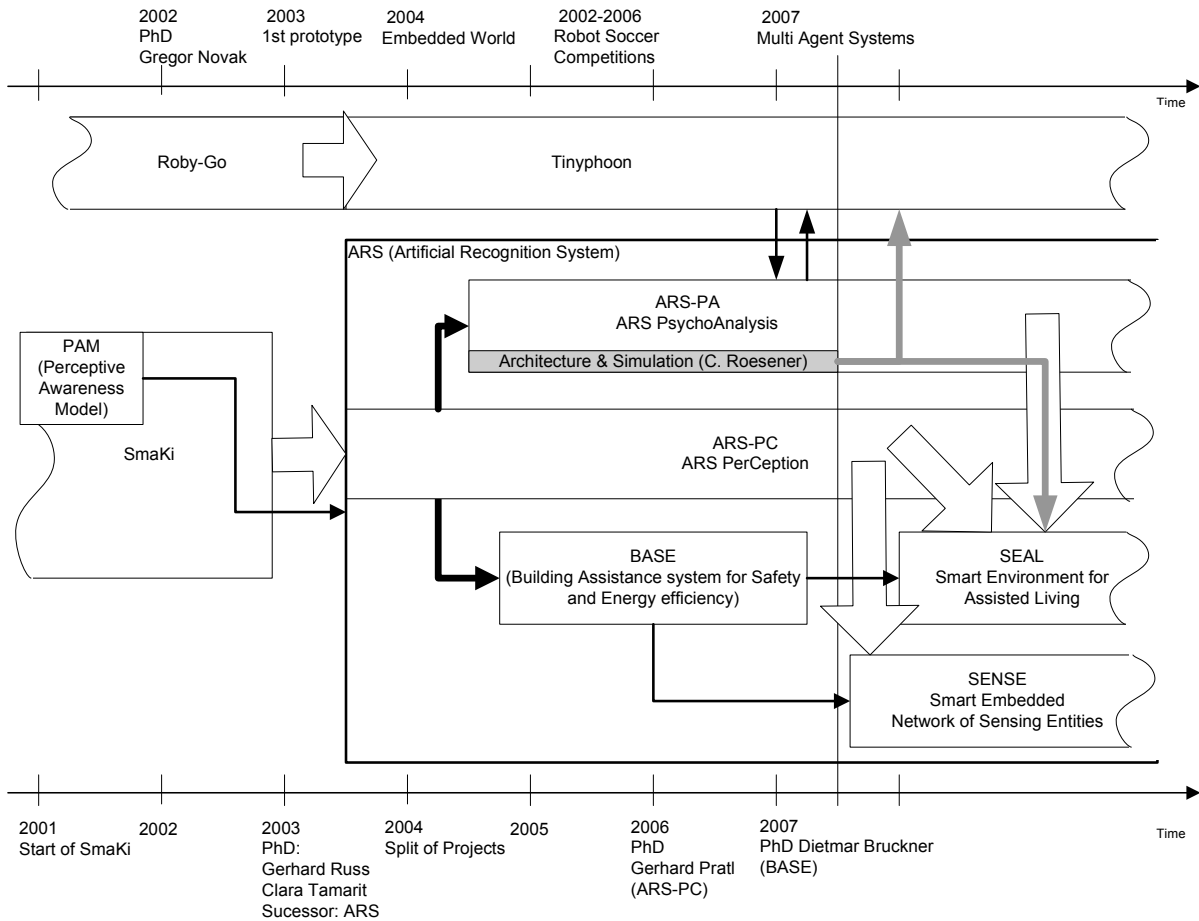
Figure 5.7: History of related and preliminary projects.

As control networks themselves seem to be generally limited in their operation range, other domains for active support seemed to be inevitable. The introduction of autonomous mobile robots was the corollary. In 2004, ARS was split into two groups of research, ARS-PerCeption (ARS-PC), which concentrates on perceptive design in control networks based on preceding endeavors, and ARS-PerceptiveAwareness (ARS-PA), which focuses on the decision making based on the symbols provided by ARS-PC (Figure 5.8). These two projects shall provide models that can be used together in one system, or as single solutions for different domains. ARS-PA makes use of theories of psychoanalysis to allow autonomous, goal-oriented behavior of robotic assistance, based on locally perceived information or information exchange in cooperation with other systems (e.g. control networks, other robots, etc.). This functional partitioning of design goals allows a specialization with a focus on a limited and more manageable group of requirements. The cooperation with other projects, e.g. the research group of Dr. G. Novak on Institute of Computer Technology, who design different generations of small-sized autonomous robots called Tinyphoon (Figure 5.7), which were originally applied to robot soccer [Novak 2005] and [Novak 2006], allowed a fruitful exchange of concepts and design methods with these research groups. As a consequence of these developments, a new control architecture for autonomous systems with focus on mobile robots has been introduced in [Roesener 2006]. The functional modules of the proposed control architecture shall provide concepts that can be used in ARS and for the robotic control of Tinyphoon or any other mobile robot requiring capabilities like autonomous behavior. The main concern in autonomous task allocation and mission competition lies

in the fact that several tasks are more of an implicit rather than explicit character, as they are generated by other demands. Therefore, a broader concept evaluating the current situation becomes necessary.
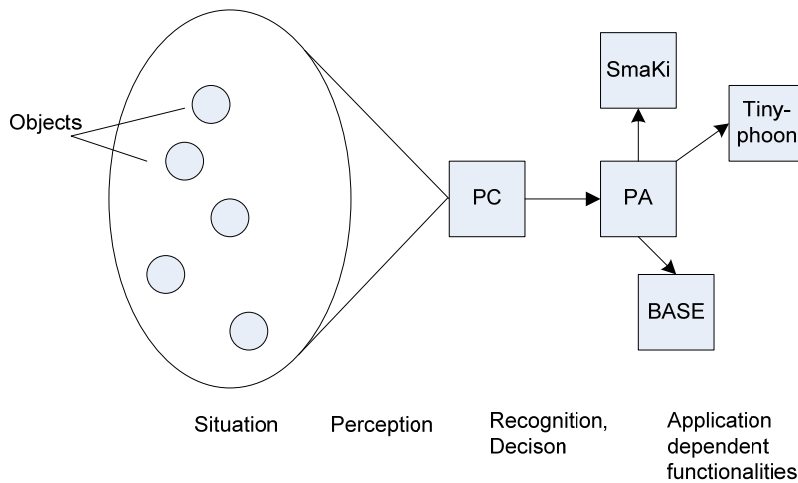

Figure 5.8: Overview of ARS projects

5.4.1 SmartKitchen

The SmartKitchen (SmaKi) project was founded in 1999 and represents the first project concerned with this topic on the Institute of Computer Technology, Vienna University of Technology. Considering technological progress in the future, as introduced in [Dietrich 2000], the intention of this project was to enhance the capabilities and functionalities of habitation based on advanced control systems, allowing safety and security, comfort, and economy on a new level. Introducing the term "Perceptive Awareness" (PA) shall emphasize that future control systems shall not simply react to given inputs, but similar to humans recognize situations and contemplate consequences. Based on studies of neuroscience, psychology, biology and other disciplines, an enhancement of the control system is capable of

− Perceiving a situation,

− Recognizing a situation,

− And selecting an appropriate action under consideration of constraints and side effects.

Based on the sensing and giving of information about situations in various form, the system allows redundancy and validity checks, retrieving proper perception results by data joining and processing. In contrast to reactive control systems, which are simply input-driven, the recognition of situations, action selection and testing of actions require additional functions, e.g. storage. The basic concept of this project is the so-called "Perceptive Awareness Model" (PAM), which is a five-layered communication model similar to ISO/OSI, with the goal to achieve a more application–oriented abstract control. An introduction to the semantics and details of this model is described in [Dietrich 2004b]. The goal was to build the first "Perceptive Awareness Automated System" (PAAS) that shall be able to prevent undesired situations, e.g. accidents. The project led to a prototype installation between 2002 and 2003 on the Vienna University of Technology, giving valuable information about further system designs.

A comprehensive overview of this topic can be found, e.g. in [Dietrich 2004b] and in two theses, [Russ 2003] and [Tamarit 2003], which were written in the course of this project.

5.4.2 ARS-PC

A direct successor to the SmaKi project is the subproject ARS-PerCeption with similar design goals of SmaKi (Figure 5.7). The task of this project is to design and implement a system capable of human-like perception derived from a variety of sensory information, and extracting essential characteristics of a situation. Due in particular to the expected massive use of sensors of different types and accuracy, traditional control designs are supposed to have serious flaws in performance, as current networks are still limited in the amount of possible inputs and outputs used for control [Kastner 2004] and [Pratl 2006, p. 1]. The model of this project is based on symbolization [Pratl 2006, p. 25], introduced by [Russ 2003, p. 71] in the SmaKi project. A symbol is an entity of information with inherent meaning[17], which can be used for control in the system. On the lowest level, symbols represent sensory information of diverse, often redundant sensor networks of different types, e.g. light barriers, temperature and pressure sensors, motion detectors, cameras, and many more. The information shall be extracted by the mechanism of symbolization, which is based on rules and sets up a context between symbol representation and the real world [Pratl 2006, p. 26]. Similar to the coding as it is used in human language, symbolization translates events, objects or any other type of knowledge into a computational "language," giving an abstract representation of the world [Pratl 2006, p. 107]. The key features of this model are:

− Control under the use of massive data from different, distributed sensor networks

− Concepts and algorithms for symbolization and processing of symbols of higher abstraction

− Associations between symbols of different abstraction levels

This project is also closely tied to the younger sister project "Building Assistance system for Safety and Energy efficiency" (BASE)

A good starting point for further details is [Pratl 2005b], which gives a good introduction and overview of this topic. Further information on this research can be found in the full description of [Pratl 2006]. In addition, information on current activities and the latest publications in this field can be found on the project homepage [eARS 2007].

5.4.3 ARS-PA

While ARS- PC focuses primarily on the perceptive capabilities of a control system, ARS-PerceptiveAwareness (ARS-PA) goes one step further. Besides observation of the environment and the system itself, decision making by choosing appropriate actions without external supervision of human operators is an important issue in modern control concepts. As the action range of actuators is comparatively limited in most control networks, service robots shall be autonomous assistants. Based on concepts of psychoanalysis, but also other areas of cognitive science, different control architectures have been proposed, which can be evaluated in a simulation. The simulation allows the use of all kinds of robotic types, starting with simple capabilities in situations of varying complexity. Due to the modular design of the simulation, different control architectures can be used to compete with each other. The described approach of this thesis is part of the project ARS-PA (Figure 5.7), proposing one potential solution for a software architecture that meets the requirements of this research project. Be-

---

[17] See symbolic representation of objects, events, and general knowledge [Pratl 2006, p.25] for further details

sides general descriptions of the overall design and functionalities of the control architecture, this thesis concentrates on the decision-making processes, and in particular on scheduling partly competitive actions of different control systems.

A general description of this project can be found in [Roesener 2006] and on the ARS project homepage [eARS 2007].

5.4.4 Future projects

There are two projects outlined for the near future. The first project is called "Smart Embedded Network of Sensing Entities" (SENSE), which succeeds BASE and parts of the ARS-PC project. The project shall improve security in public places and buildings, e.g. airports, train stations, football stadiums, public buildings like courts, etc. The system shall continuously monitor the current situation to detect potential threats and dangers. The second project is called "Smart Environment for Assisted Living" (SEAL), extending the objectives of ARS-PA by exploiting the results and mechanisms of ARS-PC and ARS-PA to build modern habitation for the elderly, aiming at relieving nursing staff by requiring as few trained staff as possible by using advanced control networks and robotic assistants.

6 System design

One part of this approach was the design of functional model for behavior arbitration, emphasizing the different control systems which shall provide a wider range of behavioral diversity, achieving complex behavior of high performance in unforeseen situations. The functional model and its derivations is used in a simulation on agents in a game-like situation. Its design is presented in [Roesener 2006] and [Lorenz 2007b], and it was developed under advice of Prof. Dietrich [Dietrich 2000], [Dietrich 2004], [Deitrich 2007] and with feedback from interdisciplinary works with psychoanalysts, namely M. Solms [Solms 1997], [Solms 2002] and E. Brainin [Brainin 2004]. The evaluation and simulation were deployed within the ARS PA project. It shall be proved that the model is capable of acting as proposed under given circumstances in a simplified environment. The control concept with emotional mechanisms shall improve the decision making of agents in general. Within the general concept, this work concentrates on the conflict resolution in action selection, which is caused by the different and partly competitive control systems.

Based on the requirements listed in Chapter 2 and the proposed methodology of Chapter 4, an overall control concept for the autonomous behavior of agents, embodied or entirely computational, is created under the consideration of the recent findings in cognitive science, as introduced in Chapter 3.1. The design is based on a behavior model containing basic functionalities of the mental apparatus as they can be found in the human mind. Its principal concept is the autonomous control of embodied agents, especially mobile service robots, facing the challenges of autonomous mission completion without predefined action plans.

Giving a brief introduction to the mental model derived from biological counterparts and its transformational into technical behavior architecture shall show the intentions and proposed achievements of novel concepts. Before describing a concrete example of behavior-based control architecture and its inner structure, it seems necessary to examine the functional relation between behavior and control (Definition 3.1). The control architecture introduced here will consist of hybrid building blocks for simple reactive behavior and blocks for deliberative (conscious behavior). Each agent has to be equipped with an inner loop using straight reactive actions, and an outer loop for reasoning (using strategies, roles…etc.).

6.1 In analogy to the human mind

As introduced in Chapter 5.3, the most fundamental essence of all human behavior is to potentiate the highest goal of an individual (human or animal), its survival. This gives a biological control system (the human mind) the virtual impulse to ensure that the organism is provided with all necessities to sustain its vital functions. However, the purpose of a technical system is not based solely on sustainability. According to our definition in Chapter 2 (2.1), the highest goal of a robot cannot be self preservation, but has to be the safety of its users, which leads to a different understanding of how to control it, as has been summarized in Chapter 3.2.2. Although aspects like reliability or availability are major factors of control, these requirements do not match with the goal of organisms, which cannot be valid for technical systems. Due to this difference in the purpose of robots, which corresponds to

further different application-dependent requirements (as defined in Chapter 2.1) which have a great impact on the behavior (Chapter 2.4), not all mechanisms described in cognitive and psychoanalytic theories can be used for the technological adaptation. This chapter shall underline the adaptation of this work by using findings of cognitive science and psychoanalysis that correspond with the different design goals.

Figure 6.1 shows the principle of perception and decision making and provides a basis for our approach. As mentioned in Chapter 5.3.2, the behavior of an individual is the result of a chain of processes and functions. Compared to Figure 5.6 of R. Plutchik's model, the model shows the filtering of perception and cognition coming from the outer world (outer ring). It passes the perceptive filter through senses (body), transforms information of different types (visual, auditive, sensual, etc.) into comparable information digits (e.g. impulses in biology, digital codes in electronic systems) that go through a qualitative filter of cognition, and sets up world knowledge (which will be stored in an extra memory system, the semantic memory. In combination with the image memory, this represents the minimum set of interpretable data).

However, the processing itself is not fixed but remains unchanged, as neither brain nor mind are completely, genetically determined or in a completely pure state. According to psychoanalysis (see also the section about psychodynamics in Freudian psychoanalysis in Chapter 5.3.3), a great deal of the described processes between the various functional units that mutually influence each other happens unconsciously. Conscious decision making, although shown as a thin surface in Figure 6.1, is a special form of processing information rather than a functional unit. It is a kind of meta-management, characterized by an explicit control of basic processes, and strongly relying on symbolic manipulation. The Ego, as defined in Freudian psychoanalysis (Chapter 5.3.3), is not identical with consciousness (*Definition 3.13*), but it gives the semantic background that refers to all mental contents which are in principle consciously accessible. Based on this idea, a memory system is introduced in the following chapter (Chapter 6.6.1).

In analogy to the emotional theories that were introduced in Chapter 5.3.1, emotions and their reactive control circuits, i.e. so-called emotional systems (as mentioned in Chapter 3.1.5, they represent competing forces as a basis for decision making and behavior selection), are represented on the conscious surface as feelings (see Chapter 3.1.6 for more detail), but they are even more than that. Emotional systems are also a set of functions which are executed unconsciously (under normal circumstances) and have a direct influence on the environment (body). This aspect of emotions shall be emphasized in our approach. Using the theories of A. Damasio and M.Solms (Chapter 3.1 and Chapter 5.3), which try to reconcile psychoanalytic concepts with recent results in neurosciences, an archetype model of emotional situation perception and evaluation has been designed (Fig. 1). This concept will be used as an example of behavior architecture to explain the features of perception and decision making.

One of the most important aspects that needs to be considered in future is to find an efficient apparatus for the reduction of the high amount of sensor data to a compressed "image," which can be interpreted and valued. This has been attempted in previous works related to this project, e.g. [Pratl 2006]. But this approach tries to go even further by using the mechanisms emphasized in psychoanalysis and cognitive science for data reduction, and by using an emotional and drive-based evaluation of environmental information. Emotions do not only directly affect action selection they can also filter and

emphasize perception itself. This happens due to their close connection to intentional processes, which can be described as a higher-level form of control.
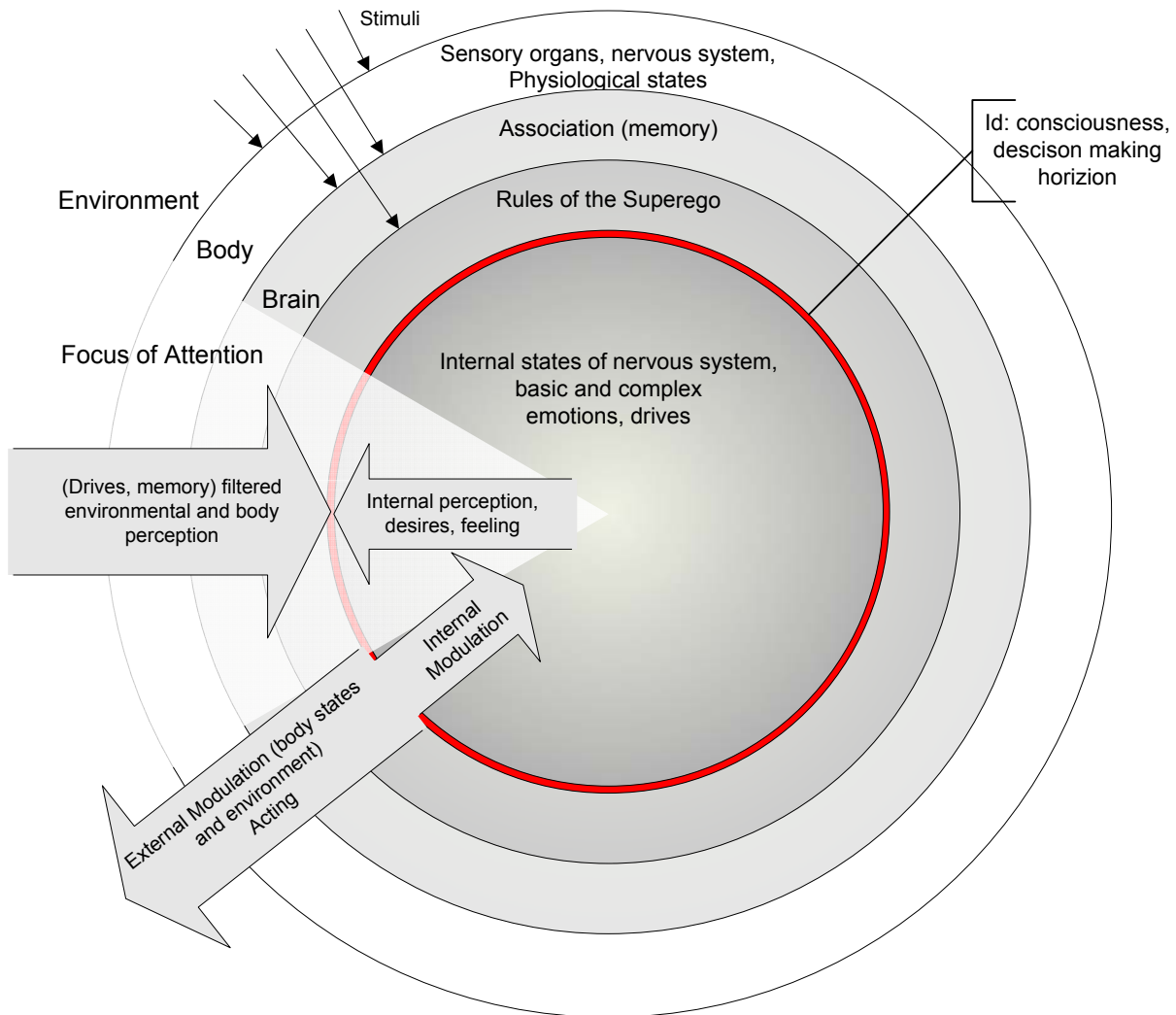

Figure 6.1: Principle of control architecture based on human archetype

The most important fact of the abstract, spherical model above is that there are two sources of information for situation evaluation, i.e. external and internal stimuli, which correspond to the drive theory in Freudian psychoanalysis as described in Chapters 3.1.4 and 5.3.3 in more detail. Internal stimuli communicate "bodily" needs. They indicate the momentary internal physiological state, as given for example by the current energy level, and the current values of in-built drives and basic (primary) emotions. Another important fact is that external and internal stimuli are filtered and modulated by various functional units, which has been inspired by a psychoanalytic view of behavioral decision making. All external stimuli from the environment are more or less influenced and filtered by memories and the superego. The superego represents socially desirable rules and blocks excessive behavior. Drives which initiate active search behavior and primary emotions, mainly fear and anger, can be compared to Freud's Id. As in Freud's theory, they are the basic source of motivational behavior. They operate on lust and aversion as the main control principles. Lust exerts an excitatory influence on behavior and aversion an inhibitory influence. The whole unit delivers a first evaluation of which class of behavior might be favorable given the present situation.

6.2 Abstract behavior arbitration model

Based on finding of psychoanalysis, respectively psychodynamics a general functional model for recognition and its action selection has been designed (Figure 6.1). The model originates on preliminary discussions in cooperation with my colleague B. Lorenz [Lorenz 2007b], and has been presented the first time to scientific community in [Roesener 2006]. After giving general conventions and about the essential roles of (abstract) images (*Definition 3.3*), (abstract) episodes (*Definition 3.7*), in recognition and action planning an universal model for behavior control has been designed. The following hierarchical model contains numerous concepts and mechanisms derived inter alia from cognitive and psychodynamic theories, and shall cover a wide range of behaviors using simple hard wired reactive and stereotyped actions, up to deliberate strategies based on action plans fulfilling needs (desires). Beside a general architecture, this thesis concentrates on the design selected functionalities, showing methods of conflict resolution and emotional evaluation which are essential functional model, while [Lorenz 2007b] focus on special methods in cognitive abilities, social intelligence, and learning.

The following chapter gives a first overview about the architecture and functional groups of the model based on the fundamental assumptions described in Chapter 3 leading to the model of Chapter 6.1. The functional modules will be described in separate chapters.

6.2.1 The worlds of a robot as a topography of the artificial mind

Just as can be found in the AI methodology, the representation of the major facts and dependencies of the world a robot is placed in is crucial for solving problems and fulfilling tasks. But before organizing this knowledge, there must be a definition of what kind of representations of the world /environment is essential for the required capabilities. The following assumptions are based on the nomenclature of psychoanalysis and has been introduced in the preliminary work [Pratl 2006, p. 52-55 ff] related to this project.

In cognitive science and psychoanalysis, the outer-world describes the physics in that the robot is embedded, but is not part of the robotic body itself. The edge of this world is the surface of the robotic body, where the interconnection with another world begins: the inner world. The inner world describes all physical parts and digital systems of the robot forming the "robotic body", which can be in different states and are a summary of the states of its systems (see Definition 3.5 of "state") and the relation to each other. The main goal of a robot is to use the outer world insofar as it supports the sustainability and balance of the inner world, which is directly linked with the inner states of the body. Furthermore, in the context of applying mission competition to a robot, I use the abilities of the robotic body, whose state is defined by the inner world, and the resources of the outer world; both goals might go along or cause conflicts. This has to stand in context with the applied tasks of an autonomous agent, that has to achieve further objectives emerging from its application and entailed requirements, which has been described in Chapter 2.1. in case of autonomous service robots in domestic use. These objectives and restrictions are variable and cannot be pre-defined without assumptions in the abilities and tasks of the embedded agent as well as considerations of the working environment and its characteristics. As this is a matter of configuration, detail examples and case studies will be described in the separate chapter of describing the simulation of this project (Chapter 7).

To find a solution to these goals, the perceived knowledge of a concept of the world, a world model, and its modalities is necessary, which will be kept in the semantic memory of the memory system

(Chapter 6.6.1). These constraints have to be configured while the model itself gives only a group of complex cause and effect principles, dependencies and probabilities in an abstract way. This abstract model itself described in this approach is now used for three different entities:

- *The recognized world*: this world gives its accurate configuration to the template of the model, containing a reduced set of constraints, which can have an effect in the near future, as their pre-conditions (e.g. the existence or non-existence of objects, etc.). It contains the inner and the outer world.

- *The desired world*: the desired world gives the optimum world and is directly linked to the balance of the inner world. The goal is that the recognized world matches the desired world as closely as possible. In the case of the desired world, the goal of the robot is to change the perceived world so far that it converges with the desired world. In the case that the desired world and the recognized world are identical, no action would be emulated. The likelihood of this situation is very low. The time horizon is long-term.

- *The expected world*: this world is the achievable world, which is a trade off of the recognized world, and has the ability to change with the given (current) capabilities of the robot. In the worst case, there is no possibility to change, which might invoke special emergency methods (= call for help). This world is of a short-term nature and is supposed to be frequently updated.

All these worlds are inevitable for the planning and solving of problems. In continuous space, these constraints can be defined graphically and mathematically as follows (Figure 6.2), defining a three dimensional action space (under consideration of robot groups) that contains a three dimensional desire field, with three major objectives:

x   The optimization of inner balances: it is a necessity to be operable in any case.

y   Accomplishment of tasks: the main purpose of the robot is to fulfill tasks, which might be given explicitly (through assignment) or implicitly (within mission completion).

z   Prosperity of the group: the operability in a group reduces the risk of being trapped as a single robot. Furthermore, the probability of successful accomplishment of given tasks increases with the size of operable robots.
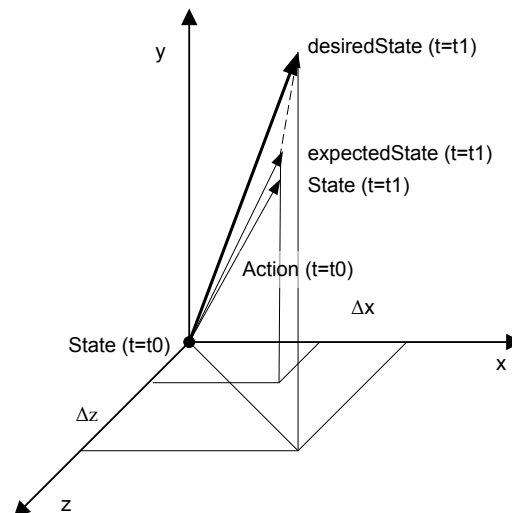

Figure 6.2: Actions in action space

The vector pointing to desired actions in general has at least one positive vector component, but in general it is expected that x, y and z are positive components. As actions, even configurable, are limited in their adaptability, it is considered that the reached (expected) state is not identical with the desired state. Due to changes during the action, even an expected state might not be reached exactly. The gap between expected state and achieved state and between desired state and expected state has to be minimized, thus solving a problem with two objectives. The mathematical decision model has the objective of achieving an expected state as closely as possible to the desired state. This form of continuous desires, are related to the idea of drives, emphasized in Chapter 3.1.4, which are continues monitoring of needs of the system, and originate desires those fulfillment can to keep the system balanced.

To simplify the ideal world, it is proposed to use instead of a continuous action space a discrete one with a limited number of possible (achievable and desirable) states, which can be generally achieved with a given sets of actions.
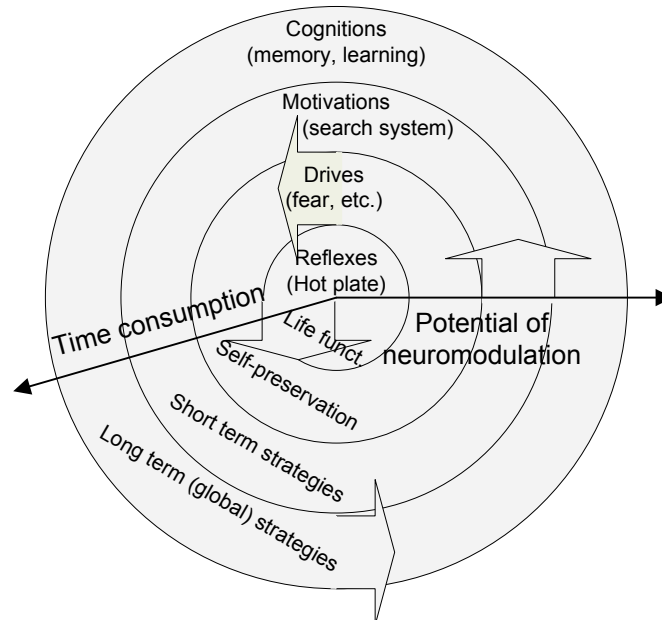

Figure 6.3: Abstraction levels of control

Figure 6.3 gives a summary of the various control loops incorporated in the human archetype, which have to be considered in this approach. They can be distinguished according to their time consumption demands and their degree of susceptibility to neuro-modulation [Fellous 1999]. Drives and primary emotions are application-dependent. Drives may be defined by their searching character. In S. Freud theory (details in Chapter 5.3.3) they are initialized by incentive stimuli. The linkage of emotions with drives, as described by J. Panksepp (whose emotion theory is described in Chapter 3.1.5) will be the basis of the evaluation system of this model and will be described in Chapter 6.2.4 and its configuration can found in the chapters − and 7.6.3. Basic needs in order to sustain and remain operable (which are fundamental requirements of the system, Chapter 2.1) any type of embodiment (either biological or robotic) requires basic needs. E.g. a sufficient and continuous supply of energy is inevitable, and therefore it is important to develop behavior which either optimizes and reduces energy consumption and on the other allows a recharging the limited resources of the body. Behavior that ends this tasks successfully can coupled with a reward, either directly with the benefit of a "healthy" system, but also

applied additional rewards, that are valid in a robotic team (society). Here, primary emotions come into play. They can be abstractly characterized insofar as they either intensify or inhibit a certain behavior in a given situation. Specific sub-systems working on this intensification-inhibition principle are be designed and described in further chapters.

## 6.2.2 Overview and functional groups of the model

Figure 6.4 shows a functional schema of the model presented in this approach, showing the functional blocks and their communication flow of the model. Similar to Figure 5.5; the interaction of an agent (human, robot or any kind of autonomous system) is a result of several functional steps, that is often in control related control architectures. The behavior and behavior selection process can be divided in the following steps.

- *Perception:* The perception module senses external stimuli originating from the environment and internal stimuli from the "body" of the robot. It symbolizes the input data so that the robot can perceive images and episodes. Finally, the symbolized content is passed on to the pre-decision module. Within this process, reflexes will be automatically executed. Reflexes are hard-wired and are not part of the mental process of a decision.

- *Pre-Decision:* The pre-decision module makes a first quick-but-rough evaluation of the current situation the robot finds itself in. Based mainly on the current intensity of the various drives, which build up the current emotional state of primary emotions like fear or anger, the perceived information is "modulated". Modulation means that certain information will be filtered and stressed. So that certain features may be stressed, others have to be suppressed. The superego, which contains social rules, can also exert some influence, as it contains social rules (or rules of the simulated game, presented in Chapter 7.3). Afterwards, in reverse, primary emotions (but no drives) can themselves be modified. Based on the modulation and the urgency, a first interpretation of the situation is executed. The pre-decision system determines action tendencies affecting the reactive control on lowest control hierarchy.

- *Decision making and conflict resolution:* Influenced by parameters, e.g. the emotion vector like urgency and complexity of the current situation, different control loops of decision making are run through, deciding on an abstract level the next steps of the individual. This provides the robot with a variety of behaviors from highly automated ones to more strategic ones.

- *Behavior:* Finally, the selected behavior is prepared and carried out in the action subsystem.

The role of emotion can be situated at several levels. Compared to former behavior models consisting only of a cognition and action module, this architecture has been enhanced by a pre-decision and a decision module. These two modules provide emotional deciding. Within the pre-decision module, the "naked" perception of the inner and outer world is filtered and modulated. This happens rather automatically and thus quickly. Primary emotions like fear or anger, but also social rules supplied by the superego are the influencing factors in this process. In turn, they can be influenced as well. Only filtered information reaches the next module for further processing.

Based on these four steps, the general functional entities shall be described in further detail. Furthermore the three competing control loops of the system will be described as form of hierarchical system, giving detailed information about its functionality and communication.

In the following chapter the functional blocks of this model shall be described in detail. The model shows certain determinism in the control on the lover control levels, described in the reactive and routine control parts, but especially the reflective part showing interplay of secondary emotions with desires and actions plans, founded on the information provided by the memory system do not follow the determinism. As also shown in the work of [Yakoh 1993] to communication and functions of control systems of higher level of complexity can hardly be reduced to a solely deterministic systems, as it often depends on non the abilities and external influences. The high parallelisms in the reflective control and its dynamics cannot be fit into serial deterministic systems, often described by state ma
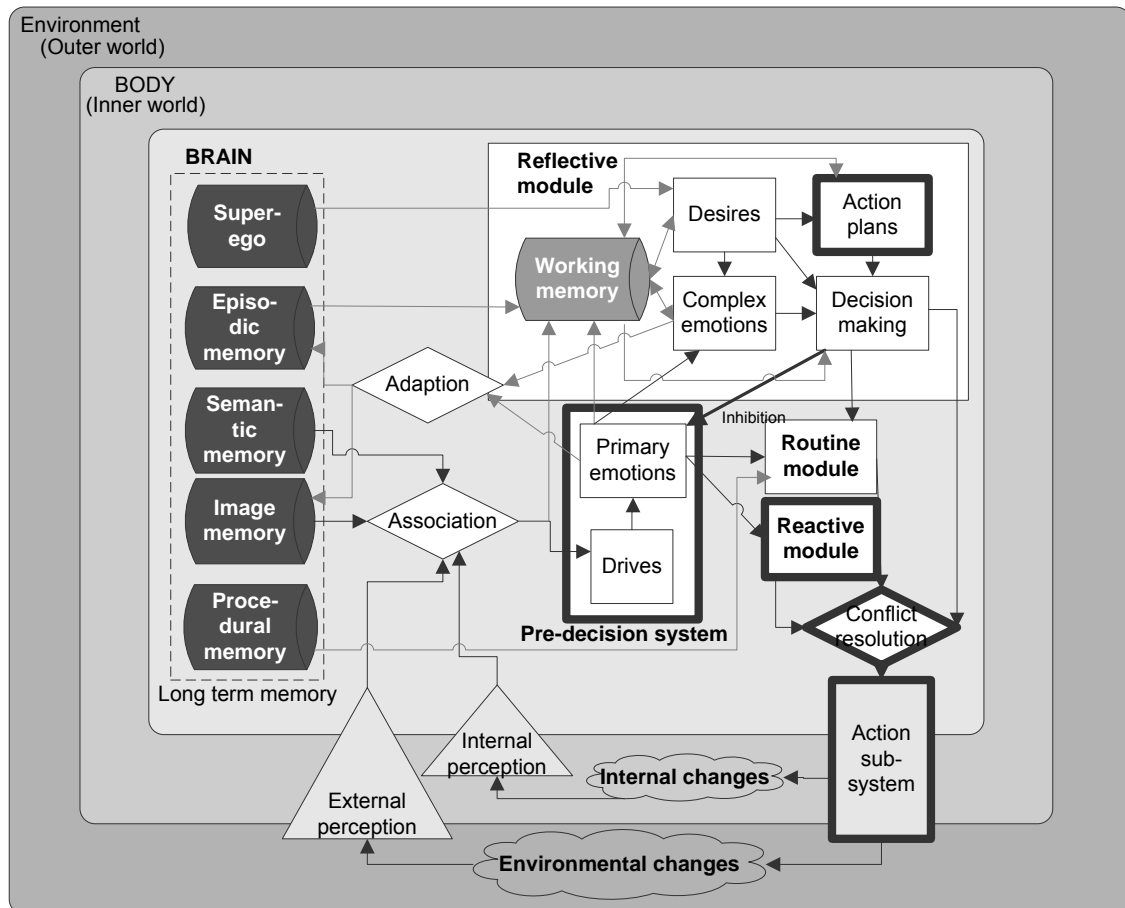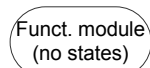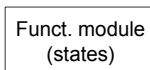
Figure 6.4: Behavior model

chines. However in the following chapters I will try to show the functional modules can be described as parallel state machines, but only under the assumption of a very limited and fixed set of memory entries, as it will be used in simulation.
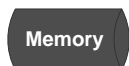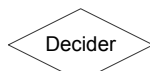
The nomenclature of the state machines will be derived form the approach of [Milner 1989] and [Milner 1999], giving a state/module combined view. In case of using this notion, the graph will have an extra mark. The nomenclature used can be defined as followed:
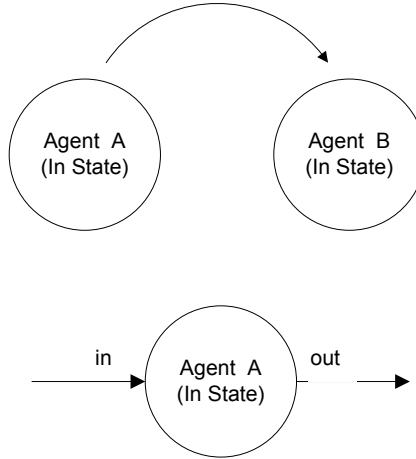
Figure 6.5: Notion of agent/state diagrams according to [Milner 1989, p. 17]

This graphical description goes along with the following mathematical description: [Milner 1989, p. 17-18]

$$A \equiv .in(x).A'(x) \tag{6.1}$$

$$A'(x) \equiv \overline{out}(x).A \tag{6.2}$$

$$A \equiv .in(x).\overline{out}(x).B(x) \tag{6.3}$$

Using the symbolism described in [Milner 1989] the diagram of Figure 6.5consists of cells, agents in a certain state, that hold data (in the simplest form it is a single value) and posses (here: two) ports. The port "in" indicates that the cell is capable to accept incoming data items (values9 and can deliver items on the "out" port. The behavior of the cell shown in Figure 6.5 can be described by the Equations 6.1 and 6.2. In analogy to [Milner 1989, p.17] the agents *A* and *A'* can either posses parameters (*A'(x)*) or not (*A'*) and the expression "*in(x).*" stands for the hand shake received at the port "in" becoming a value of x, while the agent expression $in(x).A'(x)$ of Equation 6.2 describes the behavior described by this handshake determined by the definition of *A'*. The complementary agent expression $\overline{out}(x).A$ represents the output of the value based on the definition of *A*.

6.2.3 Perception and association

Perception of this model is seen on a very high abstract level. As show in Figure 6.4, perception is a filtering process of internal and external data. This form of filter uses the method of symbolization in order to reduce the high amount of sensory data, which is expected to increase dramatically in near future and give very compressed and abstract form of information. The basic idea is, that with the introduction of symbols of an abstract entity of interpreted environmental information is proved to the control architecture that is sufficient for further decision making.

The main research in this context was done in the preliminary work of [Pratl 2006], a major work of the parallel project ARS-PC, which has been described in Chapter 5.4.2. The perception process itself will not be described in this approach. It is assumed, that the perception of this model is based on the developed methods, which are originally supposed to be used for control networks of building automation, and that adaptation for mobile service robots is feasible and can be deployed for prototyping. In future description, this approach will imply the correctness of perception and its successful interplay with the decision making model. In the following chapters there will be no further discussion about the details of this interplay and of the process of symbolization itself, except that the model can retrieve continuously symbols of the environment and robotic body containing all necessary information for further processing steps like pre-decision making and decision making.
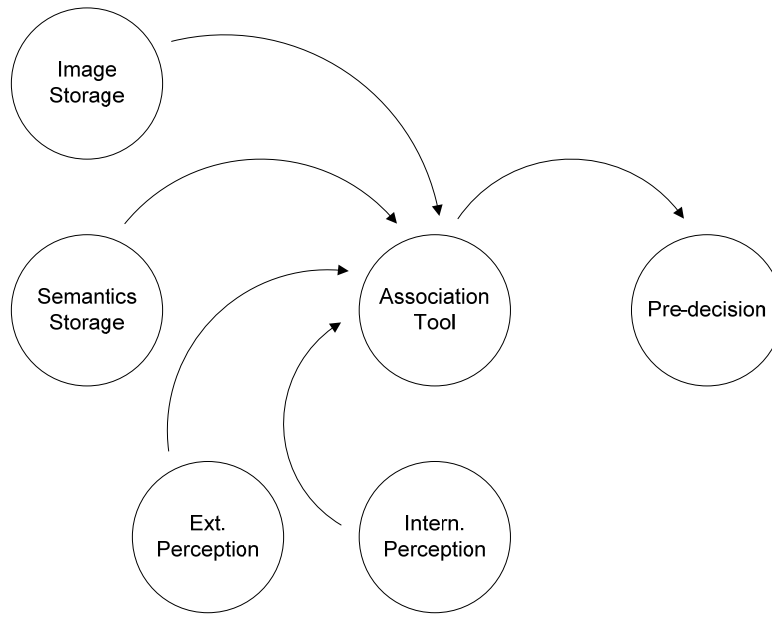
Figure 6.6: Agent/state model of association

One of the first important steps described by this model is the association of stored abstract images, (as defined in *Definition 3.3*) and the environmental, perceived current images. Their association is crucial for any further processing and decision making. The basic idea of this is, that the symbols of perception, deliver compressed environmental information which has be summarized and interpreted up to a certain content, but do not inherit an overall meaning, giving evidence to possible outcome emphasizing action-reaction relation in the lower control levels. As in psychoanalysis described perception is compressed and influenced by the memorization and experience of former situations, abstract images inherit an emotional interpretation which shall be accumulated and set in context to drives in the pre-decision module.

$$Association \equiv (in_1(perception_{in}). + in_2(perception_{ex}).Start(image) \qquad (6.4)$$

$$Start(image) \equiv \quad if\ match(image)\ the\ Finish_1(association) \qquad (6.5)$$
$$elsif\ match(semantic_{abst})\ the\ Finish_2(association)$$
$$else\ \overline{out}(cannotfind(image)).Association$$

$$Finish_1(association) \equiv (foundmatch(image_{abst}).Association$$
$$Finish_2(association) \equiv (foundsemantic(image_{abst}).Association \qquad (6.6)$$

Association module can be interpreted as a state-depend agent, that can be described mathematically according to the notion of [Milner 1989, p. 17] with the following mathematical description can be sued.

The association module tries to find the number of abstract images. Based on assigned object and values it tries to composition either via semantic rules (semantic memory), or through matching a complete abstract image. Only abstract images with a match higher than 10% shall be considered and no more than 3 images can be interlaced, forming new emotional vector by average determination of the vectors assigned to the matching abstract images. In combination with the rest factor of the for´mer emotion vector (5%), the current emotion vector can

$$\vec{E}_t = \alpha \sum_i^N w_i \vec{E}_i + \beta \vec{B}_t + \gamma \vec{E}_{t-1}, \ \alpha + \beta + \gamma = 1 \tag{6.7}$$

This method has is strongly related to the work of [Kitamura 1999 and 2002], which showing first steps of emotional evaluation, presented in Chapter 5.2.3 in detail.

6.2.4 Pre-decision

The pre-decision processing unit is a group of functional modules, which try to set the externally perceived and recognized information into the context of internal states of the agent. The agent receives a set of matching abstract images stored in the image memory, which are chosen by association module according to the filtered information of the perception units. The decider of the association module sets the extracted information of the currently perceived image in the context of stored templates in order to apply an emotional interpretation of the current image describing the situation the agent is situated in together with "experienced" (learned) or given images, which contain an existing emotional interpretation.
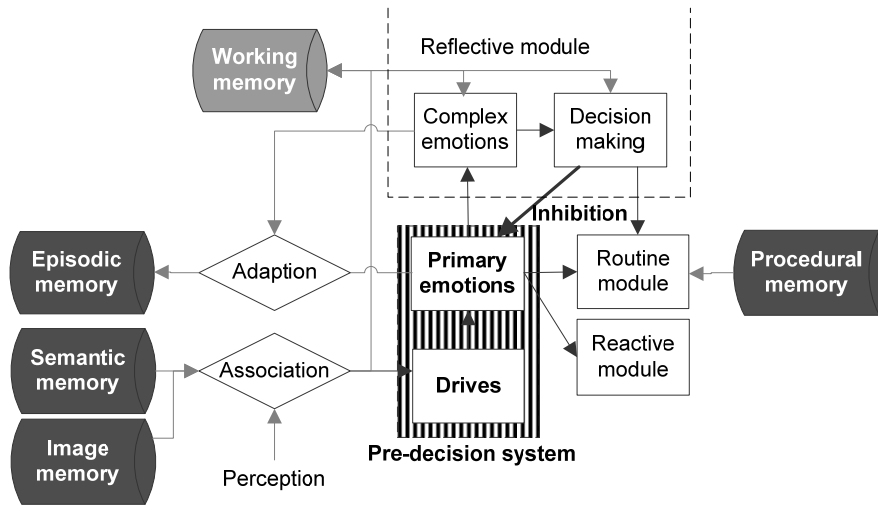
Figure 6.7: Detailed view of behavior model: pre-decision

As stated in Chapter 3.1.4, drives of organism are closely linked to the physical needs of a body, which are different for a mobile robot from an organic body. Although drives can also be made to configuration, I assume that in the near future robots will be based on electronics, and therefore the least common denominator for electronic equipment can be presumed, defining a basic set of drives for this category of systems.

Based on the assumptions of electromechanical embodiment, the drive vector of the robot contains three major components, which are considered to be crucial for the preservation of vital functions. Although the robotic body contains of different material and using different mechanisms (homoestasis, hormones, vs. semiconductors, motion primitives) compared to its biological counterpart, there is similarity. Based on general expectations of potential damage, the following 3D-vector of three explicit drives (hunger, temperature imbalance and stress) and one general value (lust), which determine the modulus of the vector, have been assigned for the pre-decision module. Using Freud's terminology as summarized in Chapter 3.1.4, the following drives of a robotic body are defined in Table 6.1. In Freudian psychoanalysis, a drive has its origin in the physical needs and imbalances of the embodiment. The number and characteristics of drives can be chosen freely, and depend on the characteristics of the body and on the environment it is situated in. However, in biology there are some common drives that are often proposed in diverse literature. In many theories, there are four crucial basic drives: hunger, thirst, temperature imbalance and sexual arousal, which have been discussed in more detail in Chapter 3.1.4. As in biological models, these drives have been applied to technical systems to represent similar physical needs in the robot. For example, hunger is the simple need for energy. There is no significant difference between hunger and thirst in the robotic body, since a robot does not retrieve energy by ingesting and transforming food into a form of usable energy, and there is no water balance which has to be considered in a mechatronic body. Therefore only hunger is considered. The sexual drives and sexual arousal used in psychoanalysis seem to be not adaptable to a unisex, electro-mechanically embodied robot without any ability of reproduction. According to the theory outlined in Chapter 3.1.4, only the ego drives for self preservation can be used in this context, which shall help to achieve the objective of the 3rd law (Chapter 2.1): "The system shall remain intact as long as possible…".

Table 6.1: Drive vector of the behavior model

Drive (values of the drive vector D)	Cause	Desire	Desire object	Regulating behavior
Hunger (H)	Lack of energy	Take energy	Energy sources of all kind	Seek, wander, escape
Temperature imbalance (I_T)	Overheating of circuits	Cool down	Cool places (ventilator)	Hide, freeze, stop
Stress (S)	Mechanical strain torsion, "pain", physical instability	Stress reduction	Team mate, safe place, dark place	Seek, slow down, hide
Lust (L)	Composed by the other three drives	Lowering the drives	Various	Various

Table 6.1 shows the vector proposed for this model. As we have seen in Chapter 3.1.4, drives are the interface between the physics of a system and its control part. Therefore, the selection of drives cannot be seen as universally fixed, but has to be determined by the individual needs of a system. The three-dimensional vector defined for a general mobile robot indicates imbalances within the robotic body which contains mechanical and electronic parts. Therefore, the last drive, lust (L), shall be a main indicator of the distance between the current state of the robot system and the balanced, neutral, "need-fulfilled" reference state.

Hunger (H) shall, similar to those in organic bodies, indicate a significant lack of energy and therefore a threat to the system's operability, while thirst has no additional purpose as electronics need only *one* form of energy and do not regulate water balance in their body.

Although not explicitly used in a number of psychoanalytic theories, in biology the importance of temperature balances in an organic body is recognized, as only within a special range optimal operability can be guaranteed. Stress (S) is in contrast to organisms an important factor in mechanical systems as there is no self-healing process implemented in the body that is based on mechatronics, where the availability and reliability of the "weakest" part determines the availability and reliability of the whole system.

Lust (L) is one of the three major forces of the psyche in Freud's theory (for a detailed description see Chapter 5.3.1).The lust drive is the sum of all ego drives that were defined above and represents a general indicator of the internal state. In terms of mathematics it is defined as an absolute value, describing the distance from the origin of the three-dimensional vector that is built up by the other three forces. Therefore, lust can be calculated as shown in the following equation:

$$L = \left| \vec{D} \right| = \sqrt{H^2 + I_T{}^2 + S^2} \qquad (6.8)$$

The derivation of the values itself are described in the Equations 7.2-7.5 in Chapter 7.6.3. Internal sensors and internal feedback allow the monitoring of internal states and can manipulate the robot's perception with a focus on external influences and a potential to redress the balance. The pre-decision system is a qualitative, dynamic filter system, which highlights stimuli that are a priority for the current needs. Furthermore, the drive vector has an influence on super-instinctual control systems. The proposed regulated behaviors tendencies are exemplary and shall not be seen to be static or fixed. Actually these are matter to configuration and have to adopted to the environment and the application field (further descriptions can be found in Chapter 7.6.3). Here the idea of desire plan will be emphasized, as they allow the execution of a complex sequence of actions in order to fulfill needs, or tasks.

Beside general assumptions in drives, the emotion vector used in pre-decision shall be considered to be fixed. Although the functionality of (primary) emotions as an evaluation system for environmental erception, there are different theories about the number and function of primary emotions, as it is described in Chapter 3.1.5. Using the four dimensional vector, including the values of seek, panic, rage and fear, shall give a first interpretation (filter) of the meaning of a current situation (transformed in a perceived image), that will be compared with experience (stored, abstract images). Emotions otherwise than drives interpret the information outside the body, and their interplay with drives, that interpret the inner needs of the body, give the balance of the robot and its environment which has to be kept or achieved by behavior.

The core interrelationship between emotions and drives can be found in the special state of the seeking system: this is an exception among the emotion systems. The seeking system, which can be also described as "curiosity" has its motivation indirectly founded in the drives, that allow to gain further information about the environment to ease bodily needs.

6.2.5 Decision making and behavior

According to [Jones 2005], behavior "is a control law that clusters a set of constraints in order to achieve and maintain a goal." In analogy to theories to of psychoanalysis behavior needs as an input data about the environment, which can be collected by sensors and be preprocessed for the control system, as well a information about, the agent itself, maintained by state information (Definition 3.4), and other behaviors, and results in commands as an output to actuators, and other behaviors. In general, we can distinguish between two different types (and hybrids forms) in robotic control [Jones 2005]:

- Reflective behavior can be seen as providing primarily "hard wired" reactions based on sensory input. The behavior represents solely the situation-dependent "reaction" to the current situation. This allows fast response due to its direct, rule-based relation between sensing and action without high computational effort necessary for extensive reasoning and the use of sophisticated world models. Although the high performance fits well to rapidly and dynamically changing environments, this type of behavior lacks the ability to adapt flexibly, and contains the risk of unsuitable responses in the long term, leading to logical traps.

- Deliberative behavior uses explicit reasoning for control. This computationally intensive form uses world models of different complexity to achieve symbolic representations of the world. The computationally intensive method is challenged to choose appropriate actions promptly. Furthermore, it might be difficult to obtain models in certain environments and situations.

Behavior control includes diverse hybrid control methods, where complex structures are built up by simple, predefined actions, including reactive components for low-level issues and deliberative components for long-term goals. However, in many behavior architectures these components are not seen hierarchically, but are built as modularized items.
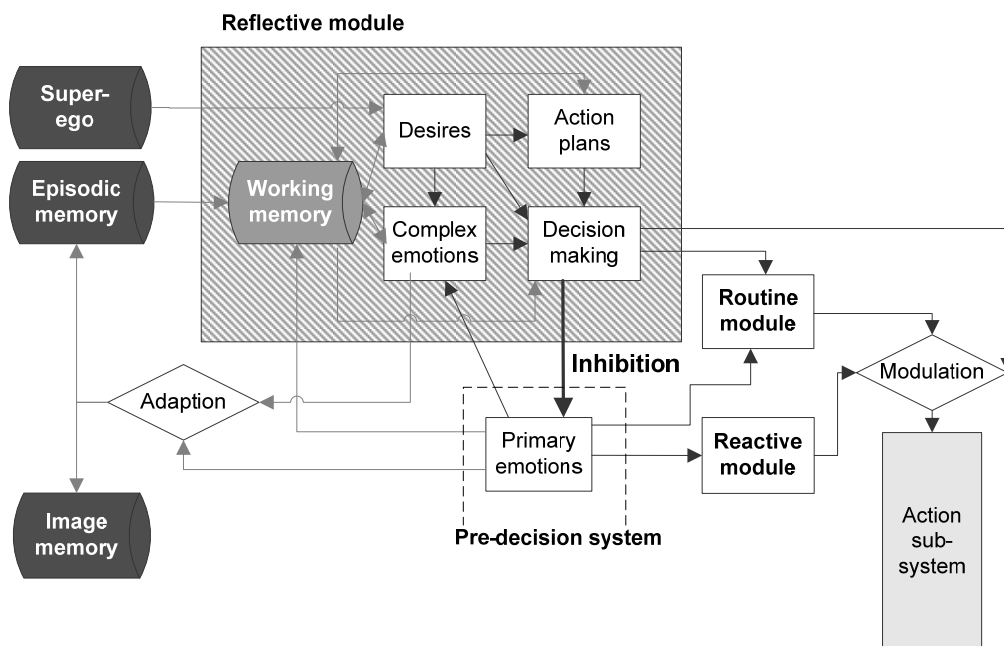

Figure 6.8: Detailed view of behavior model decision making

Before discussing the function of reflexes shall be discussed, which has been introduced in Chapter 3.1.4, other than the control systems, which will be described in the next pages, Reflexes describe mechanisms that are "hard-wired" and have no higher mental processes involved. The rules of reflexes are only driven by perception, and are not influenced by internal states (drives or emotions). They follow the stimuli-response system of Freud described in Chapter 3.1.4 and Chapter 5.3. For this mechanisms we can distinguish between two types of stimuli, that can activate behavior.

− Stimuli containing information about the organism: this can be referred to data processing (e.g. calculation, reorganizing of data storage) within the system, but it also requires additional energy of energy storage (e. g, secondary battery packs), and the activating or deactivating of cooling systems (e. g, ventilators) in technology systems.

− Stimuli directed towards the environment: in technical systems, this can be transformed into the use of all types of actors like warning lamps, sirens, the automatic shutdown of devices, etc. In case like embodying the model in mobile robots, the robotic body can be moved and changes its localization.

In a technical control system this can be employed as simple rules, which directly linked with perceived images, and therefore will be part of the image memory. Besides rule-based behavior based on the reflexes, there are three main control levels of the decision module in Figure 6.7 that can be distinguished:

Reactive processing: this is the simplest and quickest form of generating behavior. They are emotionally driven. Especially in dangerous situations, hard-wired action patterns aimed at keeping the robot from damage shall be released. Their own emotional systems represent whole control loops initiated by one of the 6 primary emotions.

Routine processing: this level executes primitive, automated behaviors. The execution is carried out unconsciously, but can be initialized consciously (through emotion or a conscious decision). The action patterns of this level are stored in a procedural memory.

Reflective processing: on this level, explicit hypothetical representations of possible alternative behaviors are used for decision making. It is here that higher-order cognitive functions and the technical equivalent of the mental model as depicted in (Figure 6.4) are located. External and internal states shall be monitored, evaluated and associated with each other. This provides the robot with a rudimentary inner model of itself, which may be referred to as "sense of self", the simplest form of consciousness.

− Having access to a memory containing former situations, their emotional rating inclusive, is indispensable for the reflective module. This kind of memory could be referred to as episodic memory. Additionally, there must also be a memory that contains semantic knowledge. In order to select the preferred action for the current situation, the robot takes into account similar previous situations, supplemented by the semantic knowledge and social rules stemming from the superego unit.

− A simple version of learning comes into play when newly experienced situations, including their emotional rating, enter the (episodic) memory, and thereby start influencing subsequent reflective processes. A more sophisticated variant of learning could be achieved via the implementation of a categorization algorithm, additionally generalizing newly experienced situations.

6.3 Behavior arbitration and action selection

Reviewing Definition 5.1, [Prescott 2007] emphasizes that "action selection is the task of resolving conflicts between competing alternatives". Originally expecting that the execution of steps in planning may automatically lead from a current state to its goals, which would be stand in analogy to theories of psychoanalysis, as described in Chapter 5.3, where the Freudian theory describes that inner state of an organism or system (*Definition 3.4*) as the source of fundamental motivation for behavior (Chapter 3.1.4). However, D. Chapman has shown in [Chapman 1985], that conjunctive goals are hard to achieve with domain-independent planning. In the 1990s, D. Chapman finally proved that there are no formally correct plan transformations achievable if actions are *not* represented as domain-dependent [Bryson 2007]. This goes along with the theory of Freudian psychoanalysis, which describes mental processes as outcome of inner mental forces, that compete with each other (Chapter 5.3), and behavior can be determined in multiple ways. However, I will show that for a design of a system, which cannot go through all steps of self-development, but requires initiation of basic knowledge, will require a minimum of categorization of action tendencies. In general, there are two major problems in action selection:

−   The determination of the available options for selection – deciding what a potential action is.

−   The determination which action shall be executed at any point in time – deciding which action is appropriate in a situation.
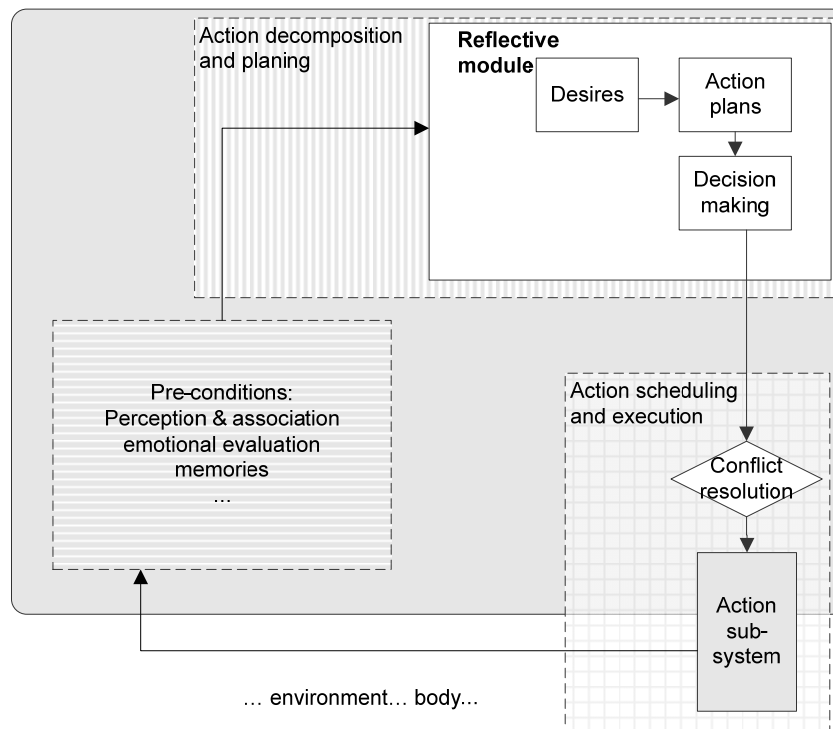
Figure 6.9: Functional steps in behavior arbitration

A planning system has to represent the world and its objects, sequences of events and the actions the agent can execute [Chapman 1985, p. 20]. [Bryson 2007] proposed the following necessary architectural attributes as inevitable to create reactive control architectures, which are capable of complex tasks.

Hierarchical discrete methods are long established programming techniques, but [Bryson 2007] challenges that these approaches lead to a more responsive and homogenous control. Standard strategies use plans (as defined in Definition 5.2), which are kinds of established sequences for action selection and might originate from experience or current processes (in this approach desired). Plans are generally sequential, giving a more or less fixed order that shall be followed in order to achieve a given goal.

Although action tendency shall be kept as flexible as possible, the learning ability shall not be emphasized in the first step and basic assumptions derived form behavioral analysis shall help to give a basic set for actions, that follow comparatively fixed rules (reactive actions), that inherit pre-defined world knowledge instead of learning mechanisms for adaptation. Therefore the memory system (besides the working memory) which used in this model can be seen as static.

6.4 Reactive control
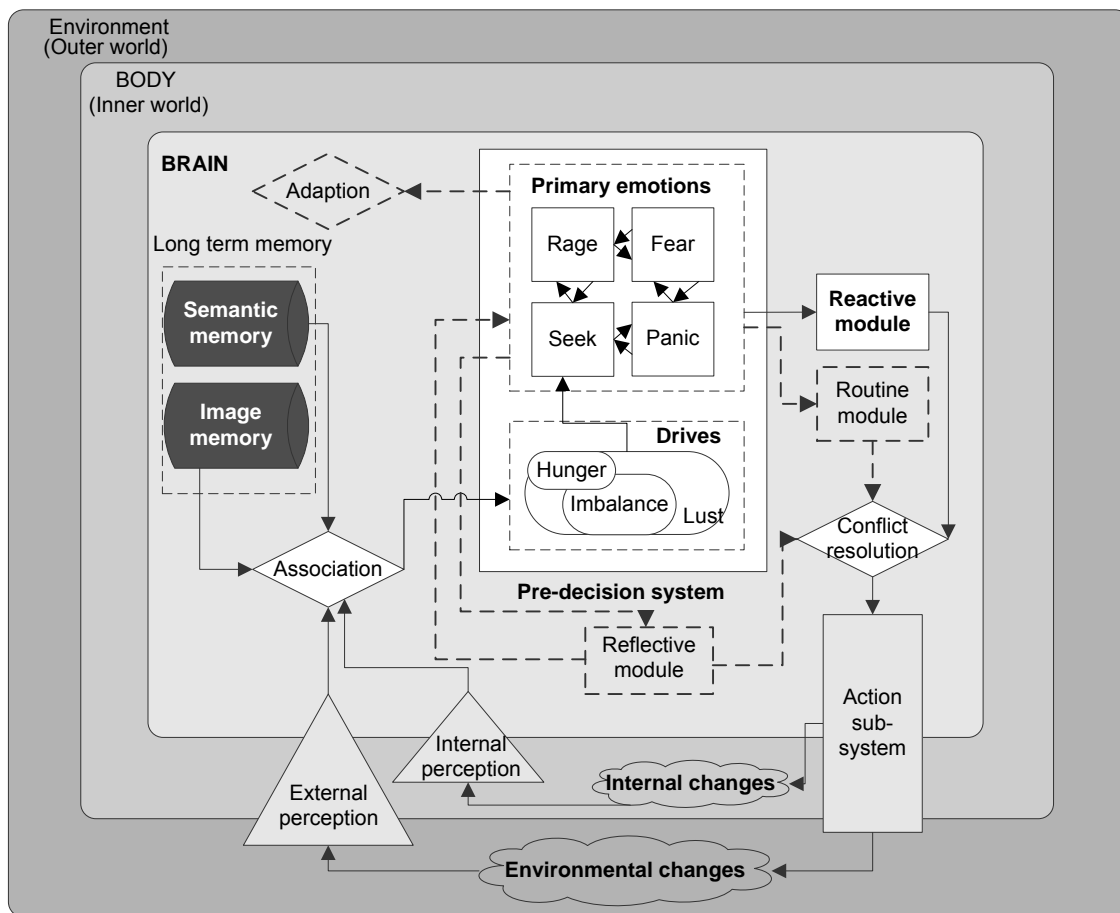
Figure 6.10: Reactive Control System

The major roles of emotions in the human mind has been systematically described in Chapter 3.1(Chapter 3.1.5ff.) and in Chapter 5.3.1 and 5.3.2, emphasizing the four basic control systems: Seeking, Panic, Rage, Fear System, as proposed by J. Panksepp. These emotionally founded control systems (presented in *Figure 3.1* and Figure 5.6) are functionally independent in the J.Panksepp's theory, possessing individual strategies and desire plans to achieve different goals. Going along with the secondary emotions (described Chapter 3.1.5), who are built up in the reflective control system, they are an origin for the non deterministic behavior described in the Reflective control. In general these sys-

tems have no hierarchical priority and keep in balance. An emotional control system can be activated by the current emotion vector to win the competition of who will have authority over the upcoming action plans in combination with the activated episodes. These internal emotional control systems are part of the reflective control. They propose competing actions and action tendencies. In this approach reactions to environmental or internal influences initiate instinctive, fixed actions (derived by the action sub system). This control represents the lowest level of the control hierarchy. There is a simple set of rules stored in the memory system (Image memory) that can be "hard-wired" and set up from the beginning. If based on the internal and external perception an abstract image can be associated matching the perceived and extracted symbols, direct rules adding simple ad-hoc actions are delivered. These It gives direct advice about the necessary motion primitives, which terminate automatically. The selected actions of reactive control are generally simple and short-term. This type of control is generally used for emergency cases, where the robotic body is in acute danger and has to employ preventing measurements immediately.

This low level control can be seen as a very classical, partly rule based approach for control. The main difference towards other approaches is its sophisticated evaluation system to set in context inner states with interpreted environmental images. The benefits of this control are:

- Immediate reaction

- Input-driven, but it requires relatively less information than more complex control systems

- Comparatively simple (rule-based) action selection

- Requires almost no memory

However, this type of control faces flaws in accuracy and adaptability. Especially in long sequences of complex actions this control cannot act sufficient. In this model this form of control shall present the lowest hierarchy allowing providing crude rule based behavior for emergency cases.

6.5 Routine control

Definition 6.1 Working definition: "Routine is a time-dependent sequence of actions (e.g. motion sequence) which is based on a more or less fixed pattern."

Routines are more or less fixed series of actions (fixed action patterns). Their pattern is formally fixed (template), but has to be adapted to a precise environmental situation (e.g. acceleration when running, velocity according to current internal state). The advantages are: fast motion, parallelism, faster decision making, freeing attention for other problems, multitasking.

In the context of computer science, routine is a section of a program that performs a particular task. Programs consist of modules, each of which containing one or more routines. The term routine is synonymous with procedure, function, and subroutine. What is known as routine in the common sense describes an action or sequence of actions, which can be activated and executed with a minimum of controlling effort.

6.5.1 Types of Routine

Routines are defined as fixed action sequences. Routines must have defined and set points where the necessity to continue or to stop can be retrieved. In general, routines can be of two types:

- Simple routines: this type terminates after executing a number of acts.

- Repeating (endless) routines: these routines repeat the sequence until the request for termination is retrieved at a set point.
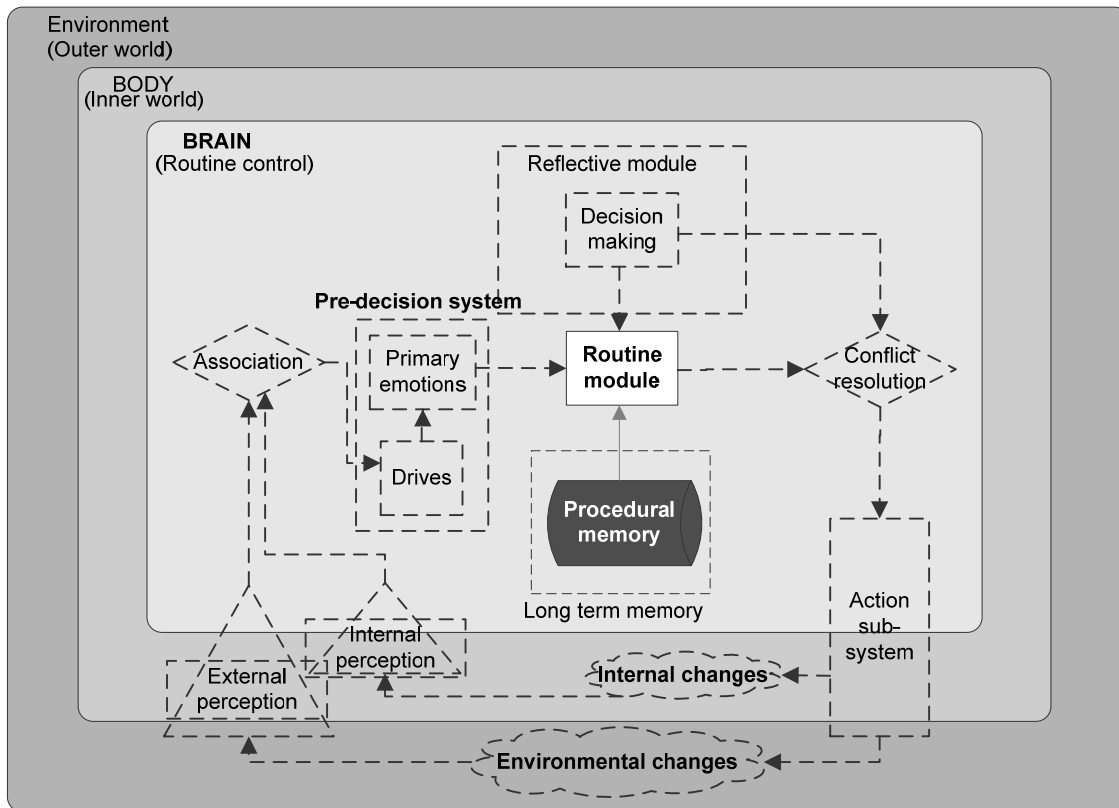
Figure 6.11: Routine Control System

In everyday life all types of routines can be found (*Figure 6.12*) Motion sequences, e.g. walking, running, sports (skating, skiing), logical complex sequences, e.g. booting a PC, staring a car, leaving the house or cooking (after recipe). The latter are terminated after the predefined action sequence by themselves, while motion primitives like running do not have a defined end.

6.5.2 Properties

Routines can be used in multitasking systems in parallel to decisions of higher levels and reactive actions of lower level. Their initialization is carried out by higher cognitive processes, but during their execution (between break points) there is no external stimuli necessary. Automatically terminating routines have to deliver a confirmation to free resources for action scheduling. Routines can be adapted to situations (attention) in "break points", which have to be triggered. The properties of a repeating (endless) routine are as follows:

- Routine needs modulation to be accurate in the current situations

- Routine can have a defined end or can be processed endlessly

- Routine has break points where it has to be triggered again and again
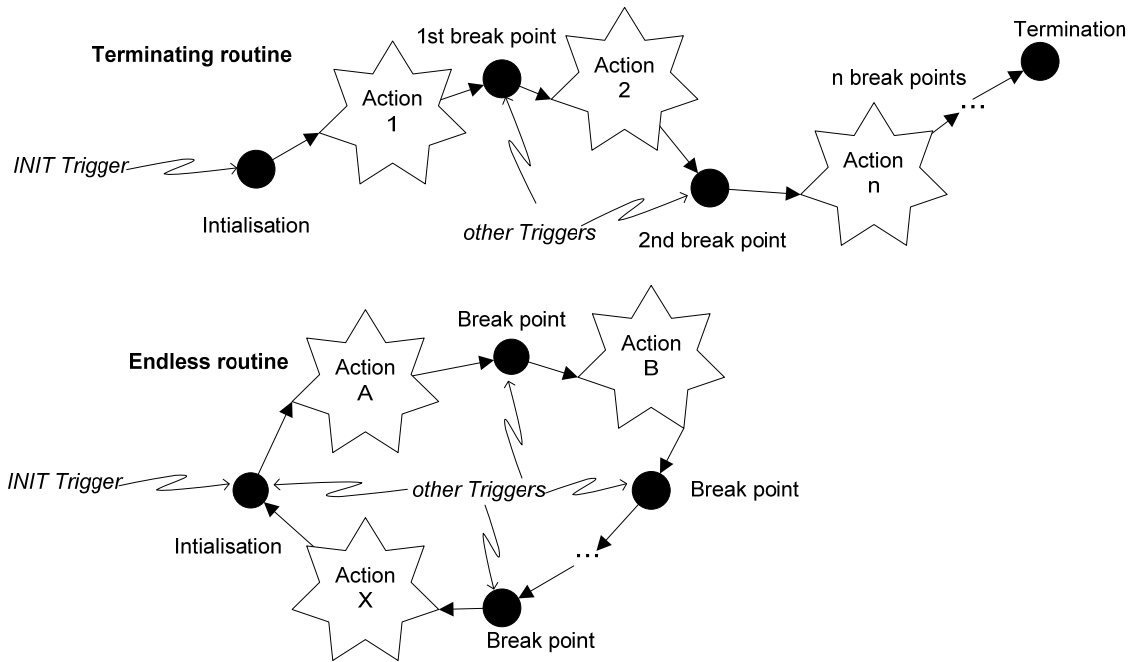
Figure 6.12: Types of routine

Legend of Figure 6.12:

Initialization=   start of routine

Action=   any sort of modular action

Break Point =   target state, min number=1

Trigger=   two possibilities: triggering of starting routine, or triggering to go to next step of routine. There
                   are to ways to design the routine: without trigger, so that in case of inhibition, the action rou-
                   tine will be completed until the next break

6.6 Reflective (deliberate) control

The last control system is the emotion vector, which is used for the sum of the stored emotion vectors of similar situations in addition to the current primary emotion vector, which contributes to the rest of the last emotion vectors by acting like an emotional memory. It is the non-deterministic part of the model containing competitive modules trying to set social/team constrains, task assignment and individual needs into context for decision making and long term behaviors.

Similar to models found in natural behavior, described in [Panksepp 1998], there are different methods to describe and evaluate behavior. The following different mechanisms are responsible for the initiation of actions:

· *Innate*: every robot possesses as set of rules, which are set to give simple action initiatives. These crude actions do not need to be learned.

· *Sympathy*: a robot can imitate behavior. If one team member shows a special action pattern, this can activate the same pattern in the other robot (flight, fight, etc.).

· *Association*: within perception, new and neutral objects can be evaluated by taking over the label for a know object, which appears in combination with the new object. This means the information content of the new event or object e_N is identical with those of event or object e_S, if they

89

$$Value(e_N) \Rightarrow Value(e_S)\{\forall e_N \mid frequency(e_N) = frequency(e_S)\} \tag{6.9}$$

· *Similarity:* if the characteristics of a new object or event are similar to an already evaluated object or event, the label can be adopted for the new one.

$$Value(e_N) \Rightarrow Value(e_S)\{\forall e_N \mid appearance(e_N) \approx apearance(e_S)\} \tag{6.10}$$

· *Specialization:* a greater group of similar objects and events can be split again.

Similarity and Specialization are methods for categorizing new objects and events in the environment. It shall help to optimize the knowledge storage for perception and evaluation. Furthermore, it supports effective action selection by categorizing actions due to their emotional evaluation.

Further description of this module will be given in the second part of this chapter describing the interplay with episodes in detail.
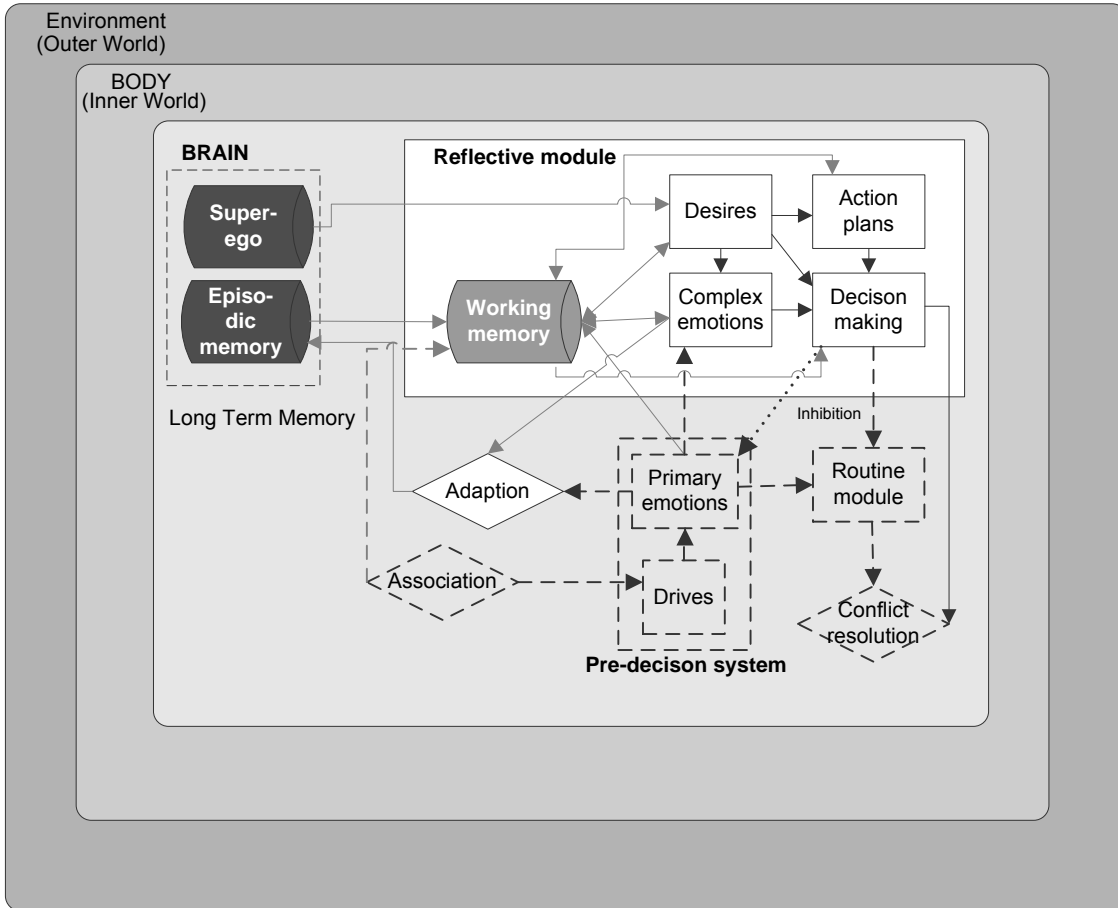
Figure 6.13: Reflective Control System

6.6.1 Memory systems of this model

As discussed in Chapter 3.1.7, there is evidence that different memories exist in the human mind. Several technical approaches, e.g. [Buller 2002] and [Dodd 2005a], have provided memory systems for behavior control. As emphasized by [Tulving 1983] and [Braddeley 1997], the separation of memories is based on the differences in information type, usage and memorizing. [Buller 2002] and [Shadbolt 2003] have proposed the following memories:

− *Procedural memory*: "How to do it?" – containing entries of behavior primitives

- *Semantic memory*: "What is it?" – containing entries of objects and instances

- *Episodic memory* (also in [Dodd 2005a]): "What can I do?" – containing entries of personal experience in this context

- *Recognition memory* ([Shadbolt 2003] only): "How to distinguish?" – containing rules for recognition

- *Working memory* ([Dodd 2005a] and [Buller 2002] only): "What happens here?" – containing accurate entries of all memories and accurate sensory information, which is valuable for on-the-spot decision making.

Congruent with the proposed structure of our approach, this model makes use of six different memory systems. In this software architecture, every functional block contains a limited-sized memory to adjust and configure the functionality appropriately for the application. Besides comparing huge memories, it holds the entire knowledge of the system. The following different types of memories have been designed for this purpose:

- *Superego*: although Superego is not a memory system in the regular sense (comparing with Chapter 3.1.7), it can be adapted as one of these. Instead of descriptions which are based on experience (which is partly overtaken by configuration) the Superego contains general social rules (which cannot be experienced), e.g. must not harm individuals, and application-specific rules (e.g. rules of the game in simulation). The memory has to be filled depending on the specific application. It is not the world knowledge, as claimed in AI, which goes along with the semantic memory of other approaches, as it is of a different abstraction level. It cannot be experienced in this system, but has to be implemented.

- *Episodic memory*: contains the connections between images and states, building so-called episodes, which are stored here. The episodic memory is important in psychoanalysis as it provides the experience of an individual (see Chapter 3.1.7). It contains abstract episodes (Definition 3.7) representing sequences of events and their potential outcome. Furthermore, it contains references to the image memory that are necessary to complete the episode.

- *Semantic memory*: gives the relation between objects, stimuli, and images. It can be designed as a world model. Its entities are no external given rules, but they are algorithms describing relations and constraints that build a functional structure of the world. This objective knowledge of the world is described in [Dodd 2005a] and [Dodd 2005b]

- *Image memory*: contains abstract images for comparison during pre-decision and reflective control. An abstract image is stored with the following attributes:

  o *Input (for comparison)*: characteristics of perception, primary emotions, drives

  o *Output (instructions)*: action directive, drives, primary emotions

- *Procedural Memory*: is used for routine control, containing fixed-time behavior patterns. It

- *Working memory:* is a temporary storage of often used and important data form the other memory systems.

Working memory

The purpose of the working memory is to give a limited set of data, whose probability of use is high, in order to reduce the processing time necessary for search and comparison in huge data storages. The scheduled data in the working memory are based on the assumptions of episode loading. The working memory of these control architectures contains the following types of data entries:

- − Episodes: the episodes temporarily stored in the working memory will be selected by preconditions due to perceptions that are given in a general situation of the robot and its environment. This may vary during simulation as accurate information is continuously collected and processed.

- − All images necessary to trigger states within episodes have to be temporarily stored in the working memory.

- − All desire planes, which might be triggered by the episodes in the working memory, have to be loaded as well. It is accepted that different scenarios might invoke the same desire plan under different circumstances. The conditions for using diverse desire plans should be a matter of learning methods in the future.

The pre-defining information can be deployed as simple, neutral data values, derived by timers, giving simple answers about the number of entities in the neighborhood, e.g. how many team mates have been passed during a journey and how much time has passed since the last one was seen. These preconditions represent a coarse estimation of the actual conditions of the environment. This means that episodes with alternative per-conditions cannot be initialized, e.g. an episode assigned to an environment with few energy sources cannot be used in parallel with episodes planned for a resource-rich environment. Unconditioned episodes are basic episodes, e.g. "get hungry," defining basic physical needs without considering environmental states. They are always initialized and must be stored permanently in the working memory. As these limited-number episodes refer to a number of basic desire plans, these have to be stored as well. These basic episodes and desire plans in the working memory form the core action space that every robot needs.

Image Memory

Although not directly depicted in psychonalaysism, A. Damaiso describes the linkage of images and drives [Damasio 1994, p. 73], as described in Chapter 3.1.1 gives evidence that images are directly linked with drives (Chapter 3.1.4)and emotions (Chapter 3.1.5). Therefore separate storage systems of images and episodes have been built up., which shall be described in detail in the reflective control (Chapter 6.6.1). The images are assigned via their ID for episodes and allow multiple assignments.

Episodic memory system for reflective control

As the model has to work out action plans (and desire plans), the observation and storage of time-consistent sequences of events are necessary. The control architecture maintains episodes as defined in *Definition 3.6* in order to predict and decide future actions and their consequences. Valuable and recognized episodes of former processes are stored in the episodic memory, and might be invoked for desire plans. Historic episodes valuable for the current situation are kept in the working memory.

In general, an episode contains two or more states, whose state transitions depend on the following criteria:
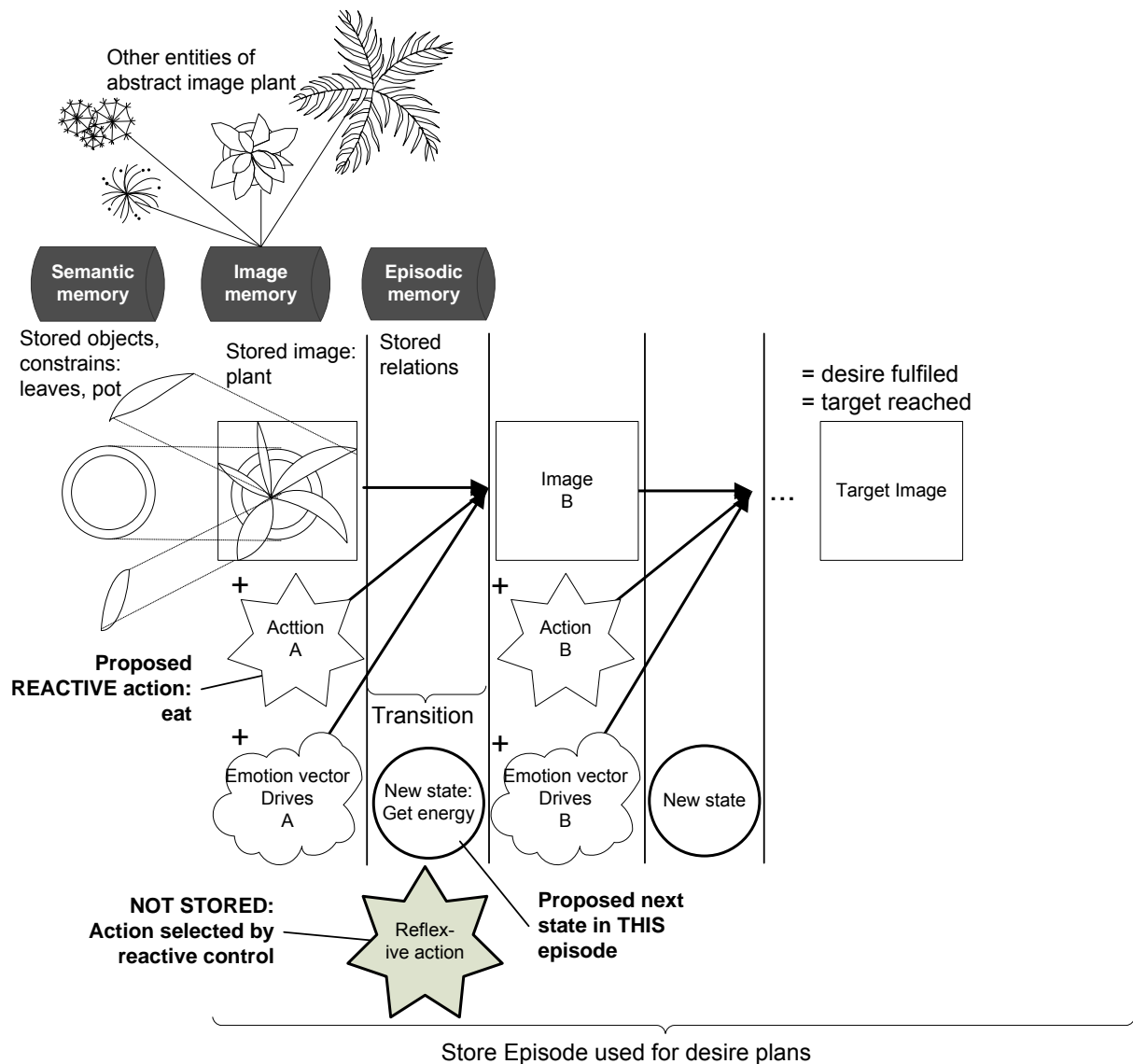
Figure 6.14: Stored structure of an episode

· Initialization: in order to initialize an episode, pre-conditions about the current environment have to be detected and matched with the pre-conditions of stored episodes in the episodic memory. Initialized episodes have to be kept in the working memory as long they their pre-conditions are valid. There is no image required for the initialization.

· The transition from (init) state to the current internal state is activated by the recognition of an image that matches the condition for transition. It is possible that more than on transition will start from a state and point to two different states in the episode. But between two identical states of the same sequence, there can be only one transition. The abstract image displaying the schema of possible conditions contains the target level, drive vector, emotion vector, and energy level.

The proposed action (or routine) for a starting situation (abstract image) action selection is not predefined by the state of the system. The transition gives a proposed action in case of necessity.

NOTE: as the action should be fixed with the situation, it does not always make sense that a new state or image activates an action. In an emergency, the fast reaction to a given situation might appear to be crucial, but the question that arises is how to achieve a long-term goal, which means that a control

93

system that considers a longer sequence of states and its transitions with their constraints has a higher probability of achieving the goal by adaptation and optimization. This previous knowledge containing information on how the state has emerged and which historic constraints have led to the current image gives the necessary information for the system's adaptation to this accurate and very specific image. As different control systems and active episodes are active, this methodology can lead to conflicts, e.g. conflicts that have been solved by selecting, scheduling and other methods of the action module.

Figure 6.14 shows the principal schema of storing an episode. The abstract image itself is stored in the image memory, a bulk of images, and their emotional representation, which can be used in various contexts for one or more episodes at the same time. These images can be coupled causing a series of state transitions, which represent an episode. The transitions and their conditions will be stored in a separate memory, the episodic memory. This entails the following boundaries:

- The same situation can be transformed into many different situations depending on the current action or internal state.

- The switch between episodes can be carried out after every transition (whenever a current situation is compared with previous situations).

The following table shows an example of potential database schemas for the episodic memory inherent with "experienced" (or set of) sequences of perceived and compared situations.

Every Situation Transition can trigger a new state and every whole episode can trigger a new state, as similar episodes can be created using the same or similar images. However these episodes have a different impact and lead to different behavior. As the images may overlap, it is likely that more than one episode might be activated and be in use. This parallelism is desired, but the conflicting action patterns have to be enforced and inhibited due to further evaluation techniques.
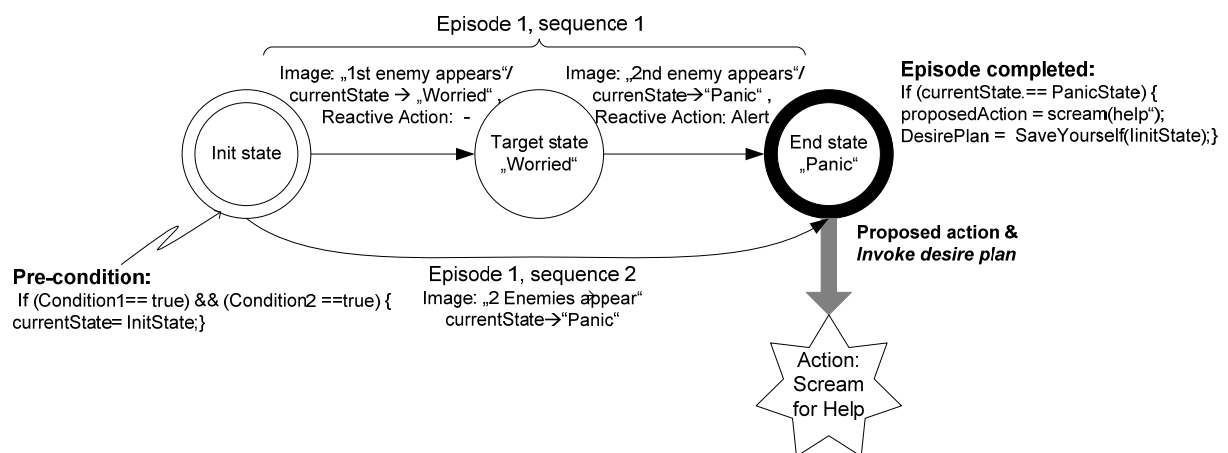
Figure 6.15: Proposed state chart of episodes

Figure 6.15 shows an exemplary episode, as it is described in the general caption (parallel of state charts) in this research. The architecture allows that the controlled embodied agent can be in more than one state at the same time. This can be described as the multiple initialization of more than one state chart, which are deterministic themselves. Each state chart represents the potential episode, that fit to the precondition for initialization. Depending on the abstract images matching with the current perceptual image, the initialized state chart can move into different states

6.6.2 Desire plans

As already mention in Chapter 6.3.3, desire plans are essential to determine the behavior to a given situation. With the increasing number of desire plans and episodes, more than one episode might match the given situation, which changes over time, as e.g. shorter, simpler episodes invoke a certain desire plan, which might be obsolete in the next state that triggers the desire plan of a more sophisticated episode. In general, active desire plans, especially in an advanced state, shall not be interrupted. However, the desire plan of a more complex episode might fit better to current events than its simpler counterpart. To manage these conflicts, the episode shall be labeled by complex emotions (emotion vector) to handle the exceptions and decide which desire plan fits best to the actual emotion vector. This may lead to an exception causing the interruption of an active desire plan. However, this kind of non-monotonic behavior means a great loss of energy as former actions become obsolete through the re-initialization of the desire plan. To minimize this loss, defined jumps or interruption respectively shall allow the change of the desire plans on a similar point.
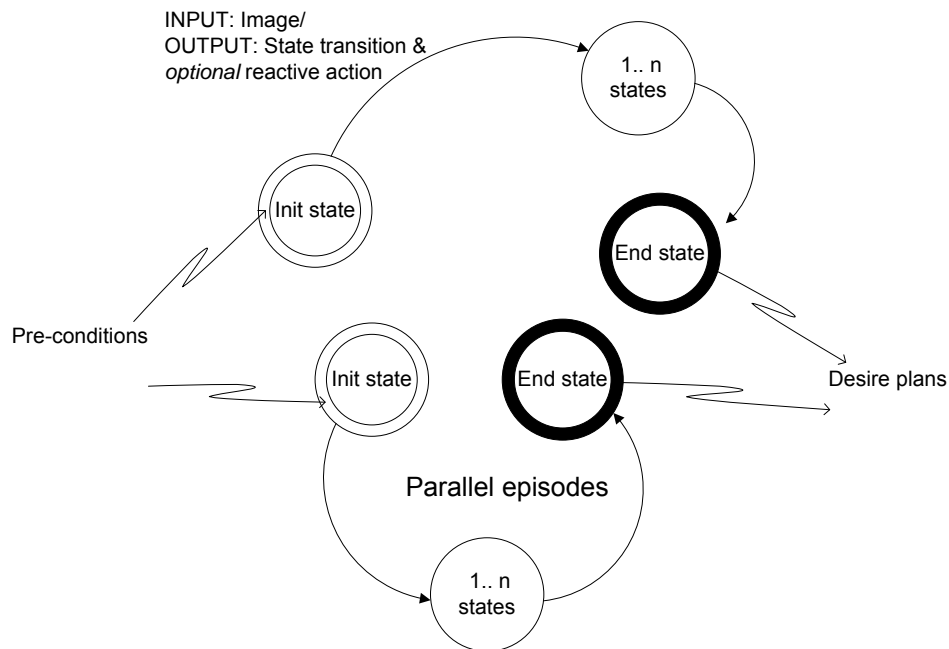
Figure 6.16: Nomenclature of parallel state charts of episodes

---

Legend of Figure 6.15 and Figure 6.16:

Init state=   First state of an episode, just depending on pre-conditions giving certain rules (sematics)

Other states=   target state, min number=1

End state=   final state invoking specific actions and desire plans

6.7 Conflict resolution in behavior arbitration

A major concern in the functional model is the competitive control systems. As different control systems propose different actions based on the different amount of information, rules and data processing, they have to be reduced and schedule resulting to one consistent executable action pattern on available

actuators. This is the major task of the action module designed for this functional model. The action module is required to schedule action sequences of competitive actions proposed by different control systems. The action decomposition can be classified as follows:

− *Spatial decomposition*: high-level goals can be split into lower-level actions.

− *Temporal decomposition:* complex behavior pattern can be seen as a sequence of simple actions.

In general actions can be executed in parallel and as a series. The proposed functional model is based on serial (e.g. routine action pattern) *and* parallel (due to its physics the embedded robot is capable to do more than one operation at the same time) actions. The nature and number of all possible executed in parallel actions depend on the capabilities of the robotic body. In general, actions can be

Figure 6.17: Interruption handling for desire plans

− Cooperative: there are actions which allow parallel processing, using different resources. Cooperative actions do not block each other's execution, and in some case even complete each other.

− Competitive: competitive actions cannot be executed at the same time, as they need the same resources and tend to achieve different outcomes.

− Mixed (combination of cooperative and competitive): in wider group it is considered, that there are actions which can bee grouped as cooperative, while others are definitely seen as competitive.

In this approach, cooperative and competitive actions can be executed. To allow parallel execution complete behavior patters have to be split into smaller entities competing for physical resources. In general, the following classifications for actions are possible:

·   Cooperative action patterns: actions which complete each other and are required to be executed in parallel.

·   Parallel combined actions: do not depend on the each other's execution, but as they do not interact and they do not affect each other when executed in parallel.

·   Sequenced action pattern: shall not be executed in parallel, but the outcome of one action is essential for the execution of the next action, and therefore have to be combined for consistent behavior.

·   Exclusive (competitive) actions: these actions stand in contrast in their intention and require the same resources for execution.

Three different control loops might propose three different actions. In general, actions will be scheduled based on a hierarchical system. Therefore the following assumptions have to be made, evaluation the importance to execute the action at the moment. Therefore a periodically priority list will be updated for scheduling, where proposed actions will be ordered based on the following criteria:

−   Sensory input: based on the requirement of special time critical and high priority data used for certain actions, priority can be shifted. This mechanism is currently not used.

−   Predefined values: certain actions have a higher priority than users, these priority can be defined due to rules; e.g. the consumption of energy the impact, etc. the result is a static hierarchical list of all actions

    (*hierarchy-pos.move > hierarchy-pos.stop,*
    *hierarchy-pos.get-energy > hierarchy-pos.communicat*e)

−   The control loop: the action selecting control loop has a hierarchy position itself, which can related to the impact these control systems are supposed to have. It is considered, that episode-selected actions of the reflective module are worth more than reactive actions of the reflective system. Therefore the (at the moment static) hierarchy: reflective>routine>reactive action is used.

−   % of selection: in case the systems does not chose one action, but more actions in parallel, the % of selection is used, and will be multiplied with all other priorities. This is not used so far, as I suppose that every control system has to choose and propose one single action at a given time.

−   Another evaluation method can be to estimate the time consumption for execution and the amount of resources required. It is considered that faster actions with fewer resources allow more flexibility in behavior. This can be used for an own hierarchy system, which can considered for additional (optional) configurations.

Due to these hierarchy systems different actions can be scheduled. However it is supposed to be possible to receive competitive actions of the same priority. Furthermore it is important to determine which actions are truly competitive and which are not. This entails additional method described in the next

paragraphs. The following example shows equations and potential action vectors for a simple robot, which is capable of executing the following motion primitives:
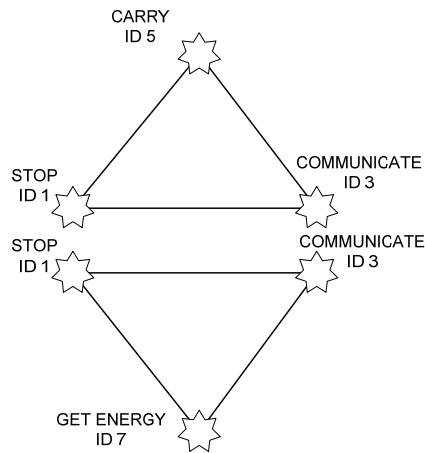
Figure 6.18: Schema of cooperative actions

− Communicate: requires resources which are not used from any other action.

− Move: is non directed changing of the local position. In the current example there is no specification about competitive directs, which have to be assumed for more complex examples.

− Stop: is equal to do no operation to change the position.

− Dance: is a special form of movement, which requires another agent.

− Carry: is to lift an entity (object or agent) in order to move it on another position.

− Get_energy: is an own action requiring special resources to retrieve energy form the environment.

− Push: move another entity (object or agent) by physical contact.
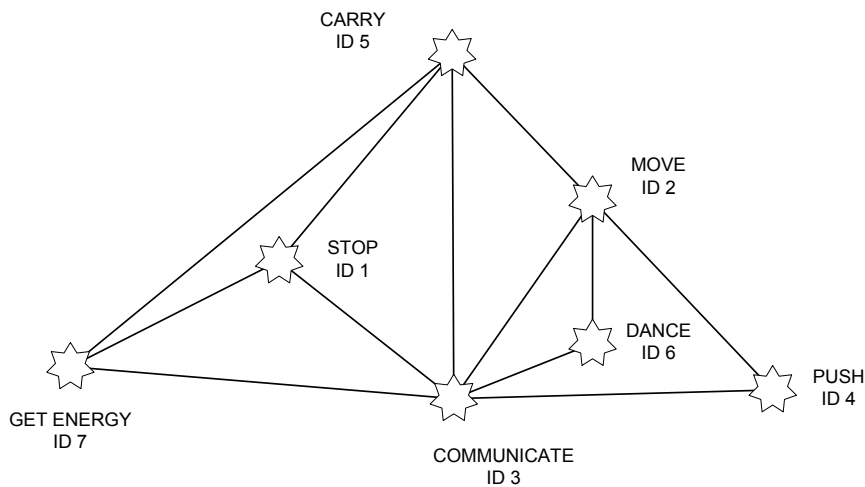
Figure 6.19: Graphical plan of cooperative actions

The robot is supposed to be capable to execute up to three of the listed actions in parallel. Figure 6.18 shows an example the graphical representation of how these actions can cooperate with or compete to one another. In the exemplary structure shown in Figure 6.18, only three actions can be derived at the same time. For example the triple *stop-communicate-carry* can be deployed by the robot simultaneously. Same with another triply, e.g. *stop-communciate–get_energy*. This is shown as a triangle connecting all three potential actions in Figure 6.18.

Figure 6.19 shows the complete graphical schema of the actions proposed for a simple robot above. Due to the capabilities of a robot, I can pre-define a group of motion primitives that represent the basic set of abstract tasks that a specific robot can execute. Building a graph of all possible triplex, assigned by the edges between (in case of Figure 6.19 7 different) behavior-entities show the action space of a robot, which has to be considered during action selection. All behavior patterns are a composition of these motion primitives. In the example of Figure 6.19, I can reduce the group to seven simple actions: 1=stop (remain at the present location), 2 = move (change location), 3 = communicate (contact with other entity), and 4 = push (change location of an object without lifting), 5 = carry (lift an object), 6 = dance (move with other entity), and 7 = get_energy (recharge inner resources). All actions need configuration data (object, direction), which shall not be discussed in detail here. Furthermore, it is assumed that no more than three basic actions can be executed at the same time. A major concern in scheduling is how to combine the actions. For every action, it is possible to define the potential cooperative actions, shown in separate tables (Table 6.2 and Table 6.3).[18]

Table 6.2: Move (ID2)

Action ID	Action ID
2	3
2	4
2	5
2	6

Table 6.3: Communicate (ID3)

Action ID	Action ID
3	1
3	2
3	4
3	5
3	6
3	7

The cooperative actions are predefined, which is illustrated in Figure 6.19. They can be either defined through into cooperation tables, or mathematically defined as symmetric cooperation matrix (Equation 6.11) to describe the cooperative working space: using the same basic actions and notion of Figure 6.19. In Equation 6.11 "0"stands for no cooperation, while "1" indicates cooperative work. The cooperation of an action with the identical action leads in this simple example to overwriting of this action (and increase of its priority). Therefore the cooperation of IDn with IDn is set "1".

Example: in case that the action *ID3* (communicate) is selected for execution, all other actionslisted in Table 6.3 (or the 3rd line /column in Equation 6.11) indicate the potential cooperative actions. But in case *ID3+ID2* have been selected, what is a possible third action?

The following rules are derived based on the graphical and mathematical description: in the actual case, only actions, that have "1" in the second and third column can be chosen: e.g. *3+2+7* cannot be derived, as *3+2* and *3+7* exists, but *2+7* does not exist ("0"). Therefore all feasible combinations (on

---

[18] Please note that the cooperative table can be used for all connections (only views of the table needed).

basis of the combination 3+2) are as follows: *3+2+4, 3+2+5, 3+2+6*. Graphically, this means that every triangle of graphic on the plan of Figure 6.19 is a possible combination in cooperative behavior. (This system can also be transported into a 3D model so that you can allow 4 actions in parallel. Mathematically, it is the same problem and can be traced back to the 2D problem.).

$$
\overline{Act_{coop}} =
\begin{matrix}
ID & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
2 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\
3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
5 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
6 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
7 & 1 & 0 & 1 & 0 & 1 & 0 & 1
\end{matrix}
=
\begin{pmatrix}
1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\tag{6.11}
$$

Competitive action pattern

The competitive action decision is used in the case of a cooperation test failure. Using, the example of above the version *3+2+7* was not possible, but the combinations of Equation 6.12 and 6.13 can be executed:

(3+7) OR (2) (6.12)

(3+2) OR (7) (6.13)

Here, the three-action problem has been changed into a competition between four action combinations (patterns), shown in Equation 6.14:

(3+7) OR (2) OR (3+2) OR (7) (6.14)

Among the possible combinations there are 2 possible action pairs and 2 single actions are capable to meet the requirements. As the pairs are supposed to have higher impact and flexibility (*3+7*) OR (*7*) and (*3+2*) OR (*2*) can be reduced to (*3+7*) OR (*3+2*). However this example has not considered the priorities.

Timing

The problem of timing and adaptation has been discussed in Chapter 6.2.1, the adaptation and timing of behavior is essential. The different control systems do not work synchronously, proposing different actions which shall be executed in the different execution time. To achieve a consistent behavior, the following set of rules shall guarantee a more stable and constant behavior:

- Generally a current action cannot be interrupted (except on a clear break-point), to ensure the system do not end in an instable state.

- A set clear break-point indicates a safe way to end an action before termination, and can be used for event handling.

- A special warning requires emergency stop is the only exception that allows termination before time. It shall not be used in regular operation.

Besides the simple examples, using atomic and fixed basic actions, actions can be adapted according to the changes in the situation (Move: move faster, move slower, turn left/right,…) An action pattern must be capable of blocking short-term action for a sufficient amount of time in order to reach its target in more than one situation. This inhibition function can have three different domains for execution:

- Block action until the break point of executed action.

- Block action until executed action ends.

- Block action until executed action leads to a new situation.

Example: as long as action 2+3 is active, action 7 is blocked. As soon as one of the combined actions is terminated, action 7 can be imitated in the case that the remaining active action is compatible. Otherwise, another action of the next highest priority is chosen. An active action (as long as it is not a routine at a break point) shall not be terminated before its scheduled end as long as there is no emergency value overruling the scheduling. Actions contain timers, which are directly linked with their priorities.

7 Simulation and visualization

This chapter presents a comprehensive discussion about the purpose and structure of simulation in this research. As outlined in Chapter 2.5, it is necessary to apply the abstract model to a concrete situation and evaluate its functionality in order to emphasize the consistency of the approach, but also to stimulate further research. Designing a universal concept does in itself lack the application-specific configuration, which is a matter of optimization. The design of this model has 3 main areas of configuration:

− Environmental configuration: all types of perceptual memories (e.g. image memory, semantic memory, episodic memory etc. as described in Chapter 6.6.1) and the Superego, which contains pre-defined rules (Chapter 6.6.1), can be fed with an appropriate set of data to sufficiently describe the environment in which the agent (robot) is placed. This kind of application-specific semantic knowledge about the environment comprises a combination of given rules and important key facts to facilitate the agent with all necessary aspects to find and achieve an optimum, but not pre-defined behavior in fulfilling tasks due to the application. This environment, real or simulated, shall give a basic understanding of interaction. In the case of testing environments, it is important to cause abstract situations, so that a wider range of special cases can be covered, but it has to be kept inclusive due to the challenges autonomous agents have to deal with in general.

− Robot configuration: the action module requires a set of basic (abstract) behaviors, describing the robot-specific capabilities based on its physical structure. Even in the case of a simulation which uses a highly abstract agent, a minimum set of assumptions about capabilities have to be adopted. This configuration can be in the context of the basic requirements of the physical shape of a robot that is designed for domestic application, as defined in Chapter 2.3 and emphasizing reliability and maintainability in processes in which people without specific training are involved. But it can be also designed in a completely abstract sense, emphasizing the basic capabilities by deploying the simplest form of behavior arbitration.

− Model configuration: although the number and interplay of functional blocks described in this model shall remain unchanged, there are degrees of freedom in the adaptation and configuration of certain blocks. However, in particular the reactive control is tightly coupled with the purpose of the system. Therefore, desire plans are directly, and the number and specification of secondary emotions are indirectly coupled with the application of the system control based on the model.

In contrast to simulation tools concentrating on one single application, using a specialized robot cannot fit the requirements of a universal concept, as optimization and specialization have a strong influence on the design, using presumptions which cannot hold when being adapted to another application. Similar to code reuse in software architectures, a modularized simulation and control architecture shall enable the reuse of modularized functionalities and enable adaptations.

The current simulation contains a specific configuration based on M. Toda's idea [Toda 1982] and described in Chapter 5.3.4, which was used for emphasizing the basic potentials of the human mind, which were the objective of this model as well. A game-like situation considering a sense of competi-

tion emphasizes the qualitative aspect, as on this abstract level a quantitative analysis might not produce the necessary significance, and a quantitative testability is still questionable, as has been discussed in Chapter 7.7. The final chapter gives a short overview of the results.

7.1 Methodology and purpose of simulation

"*...what finally counts are theories and ideas, no matter where they were originally hatched, either in an armchair or in an experiment. If an idea is good, it will eventually find a way to be experimentally tested, while a blind experiment produces only a trickle of possible facts out of the whole ocean of possibly observable facts*" [Toda 1982, p. XIV].

The idea is that the predicted optimum behavior of a human in a given situation is identical with the proposed behavior of the software architecture (Chapter 4.2, the Turing test adapted for emotional intelligence). To state it more simply, a robot shall behave as if it would be directly controlled by a human operator (which means that the human operator has the same restrictions in behavior and perception as the robot itself), making the same decisions although the behavior is solely determined by the behavior model.

The major question in this context is how to evaluate this prediction and how to compare it? A direct comparison of human operating behavior with the software architecture can be deployable in simulation, but is an individual person representative for mankind in general? A qualitative statement is hardly feasible at that level. The results depend strongly on the experience, environment, application, etc., which do not necessarily match. The setting and configuration of the experiment can turn out to be inappropriate or lead into unintentional, unpredictable operational sequences. The individual operator of a robot is a representative of a skilled profession and requires skills that not every individual has.

A second attempt would be to compare them in a different manner, trying to evaluate the question whether a human accepts a robot as "humanoid" assistant? This question requires psychological background and has a high impact due to side effects like human-machine interfaces, communication skills, etc. The proposed model that is embedded in robots has so far not emphasized human-machine interfaces. However, in this question the appearance and communication with the human colleague is crucial for the rating of its behavior. The estimation of "intelligence" requires a minimum of communication. Even in the form of silent observation, there must be a minimum of tractability to understand the behavior of the robot. This still does not answer the question whether the behavior is truly "intelligent", or whether it just so happened that the robot accidentally achieved the apparently optimum solution in the given situation. The reason for the chosen behavior is always transparent. The repeatability might give some evidence of the inner semantics but this is not satisfactory and does not necessarily guarantee the predictability of future behavior. Furthermore, the repeatability is an abstract property and is often not feasible in real world situations.

In this approach simulation shall help to emphasize the inner processes and give evidence at least on an evolutionary, animal like level, showing that basic mental abilities can be emulated by the presented model of Chapter 6. Based assumptions, that are summarized in Chapter 5.3.4, a similar attempt like M. Toda's one-person-game is used for evaluation, following the same principle idea, i.e. to create a simplified hypothetical environment, which still contains the essential characteristics of the human

environment and provides a set of conditions that challenge the subject as a problem-solving system. The situations vary in complexity, but always focus on the provision of constrained problems evoking various conflicts for qualitative evaluation. This simplified environment shall provide the basis for the initial situation, which shall be enhanced step by step to a more sophisticated environment which becomes closer to the realistic situation in the human living space.

To validate a control system is extremely complex without sufficient simulation [Mondada 1994]. The interaction of an autonomous robot with its real world environment is highly dependable of the physics of the robot and the characteristics of the environment. The control model provided here shall give a coherent and abstract basis, which is robot independent. This design goal stands in contrast to most robot control programs that ran solely on single type of robot or a limited group of similar size, shape and habits. The optimization of the control systems entails, that the "controller" cannot or can hardly be separated from the rest of the system. This cannot fit the idea of a universal concept for high-level design, which might be obfuscated by optimization. Simulation shall assist in testifying, what is often inadequate with real robots: giving a comprehensive overview about the interaction between embodied agent and its environment, but even more about interactions between functional entities within control side and inter-process communication. Based on the dimensions require

As proposed in the concept formulation (Chapter 2), the requirements of a robot can be divers depending on its application. This entailed the requirements of an abstract architectural design (described in Chapter 2.5 which has to be taken over in the simulation design. Although, a specific application was kept in mind in design of this system (Chapter 1.1) the methodology proposed for this approach was to give universal design that shall avoid limitations through specialization and optimization for a single application field (Chapter 4.2 and Chapter 4.3). The simulation part of this project shall give space to run the model, which shall help to understand which mechanisms can lead to successful control architectures. It shall help to configure and emphasize this abstract model for wide range of applications. In the first step the simulation environment cannot provide all properties of the real world, but using the idea of M. Toda (explained in previous Chapter 5.3.1 and in the introduction of Chapter 7), which emphasized an artificial environment, which is better to manage, but still contains enough potential to provoke complex behavior. Therefore agents (Definition 3.32) of higher abstraction level are used, that shall emerge behavior due to model based control architecture. Similar to the requirements on the control design, emphasized in Chapter 2.5), which focuses on the portability, modularity, scalability and reusability of the architecture, the simulation must fulfill following design goals:

- Object-oriented programming paradigm

- The same architecture (programmed) shall control different kinds of robots

- Exploration of different paradigms, configurations and methodologies

- Extendible in application and robot-environment interaction

- Interface for extensive visualization

Chapter 5.2.5 shows the difficulties of optimized and robot specific simulation. This chapter shall show the architectures used for simulation and emphasize its benefits.

7.2 Simulation structure

The idea of this research is to show that pre-defined and manually optimized behavior of an agent (as summarized in Chapter 3.3.2) cybernetic robot can be reached with the fundamental assumptions of psychodynamics and the latest theories in neuro-psychology describing the basic capabilities of the human mind. The simulation environment shall provide a simplified environment for the robot, which is equipped with the behavior-control software. For reasons of scalability and adaptability, the software of the simulation experiments is divided into three major parts (Figure 7.1):
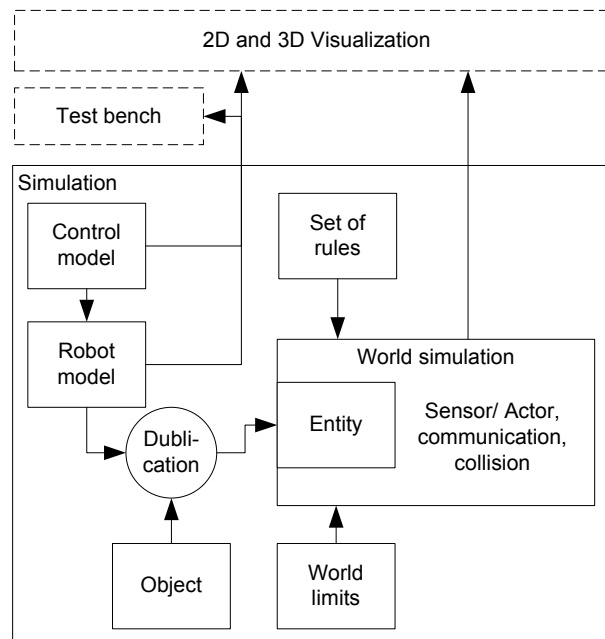
Figure 7.1: Simulation engine

- Behavior simulation: various modules of the general behavior arbitration architecture have to be adapted to suit a specific environment. The tasks the robots need to fulfill require a specific set of behaviors. Moreover, the selectable action patterns need to be adapted to the "physical" capabilities of the robot. The selection of an action pattern is based on the emotional interpretation of incoming data. Therefore, the robot needs to be equipped with a suitable set of drives and primary emotions.

- Robot simulation: the robot simulation represents the missing link between the simulation of the continually changing environment and the behavior simulation. The embodiment of the agent as a robot in the environment gives the control architecture the tool to act and perceive in the (simulated) environment The selected action patterns have to be transformed into executable commands. This transformation is robot-specific and depends on the capabilities and physical structure of the simulated robot. In the performed simulations, an abstract model of the Tinyphoon robot was chosen.

- Environmental simulation: the environment consists of sets of entities (robots) which can influence a variety of movable and non-movable objects. This is established according to given physical rules.

Separating the simulation of the environment from the simulation of the behavior of the robots and the simulation of their physical characteristics allows the easy variation of the simulation experiments.

Figure 7.2. Overview of simulation concept of environment simulation

106

The simulation experiments are conducted with Anylogic, a Java-based simulation framework, which has been introduced briefly in Chapter 5.2.5. The language Java was chosen because of its universality, speed, and simplicity of implementation. Similar to other commonly used simulation tools Anylogic provides abstract libraries for agent-based simulations that support and improve the simulation experiments of the behavior model. The three main parts of the simulation software (environment, robot, and behavior simulation) are set up within different libraries. This allows e.g. that the action selection units can be changed independently from the environment simulation. This structure facilitates the "wrapping" of the simulation and makes it possible to work with abstract action patterns rather than proprietary machine instructions. As a consequence, the behavior module can be simulated without the necessity of having a detailed knowledge of the functioning of the moving unit of the robot.

Figure 7.2 shows the class diagram and signal handling for the use case of positioning and collision detection in an active entity (which is an agent in the current setting described in the class model of Figure 7.3), containing configure the necessary configuration for the simulation engine. Defining a radius of sensing and action, as it will be necessary for defining and distinguishing scenarios as described later in Chapter 7.7.1. The active entity can collide with other entities, which can be either passive (energy sources) or active (other agents, which might be a team mate or enemy), or the entity can reach the world limits, which it cannot trespass. There are in general two types of collisions: physical which describes the touch of the robotic body with the environment (object, other agent etc.) or the collision of action space, which represents the change in the action space, where all objects can be sensed and used by the entity. These and other use cases described in UML have been used to give the generals settings of the simulation engine in Anylogic which refer also to the semantic rules about the simulated world, that are used for testing in of control architecture (further description can be found in the following Chapter 7.4).

Besides the simulation concept, as described in Figure 7.2designed in AnyLogic, the visualization of the object has been derived with in the Project ARS-PA, giving a observable representation of the current simulated situation in the world [Deutsch 2007]. Otherwise than the test bench, which shall help concentrating on the inner states and functions of the control architecture, the 2D (and 3D) simulation shows the full concept emphasizing the interplay and behavior of the agents and other entities. The simulator is a distributed system and has been described in [Deutsch 2006], [Roesener 2006] and [Deutsch 2007] in detail.

7.3   The Bubble Family Game (BFG)

Inspired by the idea of M. Toda (described in Chapter 5.3.4), the performed simulations are set in an environment similar to that described for fungus eater of M. Toda. A game like environment has been designed for simulation, which shall represent a simplified working environment, which still inherits conflicts of higher level, challenging control systems and allow the qualitative measurement between competitive groups. The idea of this game is to create an evolutionary team of agents, set in a scenario of clear set of conflicts and constraints, which are manageable, but still lead a wide range of overt behavior. The embedded agents (bubbles) are grouped in teams (of various sizes, but in most cases teams of 3 or 5 robots), that try to compete and survive in an environment, providing energy resources, that allow to survive in this environment. Depending on the scenario, there can be additional tasks,

like foraging, collecting objects, for the benefit of the group. Additional tasks like this shall emphasize the conflict of primarily bodily founded needs (referring to the idea of drives and primary emotions, as they have been described in 3.1) and a higher purpose (which can be seen as a kind of Superego in Freudian psychoanalysis, Chapter 5.3.3) of the robot. The main dilemma is that all higher purpose, cannot be implicitly rewarded, as it has no direct benefit for the robot and its sustainability, and might cause a dilemma: as physical inability due to maltreatment of a robotic body can lead to the inability of fulfilling *any* task. This has to be considered and emulated in simulation. To invoke these problem constructions, no complicated set of rules and role management seems to be necessary. Looking, at M. Toda's theory, described in Chapter 5.3.4, there simple foraging tasks in combination with constrains and limitations in the working environment can cause sufficient level of complexity resulting in observable differences in behavioral patterns, similar to evolutionary biology, observing animal populations.

Due to these design goals, a set of rules of the game and constraints describing the environment have to be set up and defined. These rules and other parts of configuration will be discussed in detail in the following chapters, giving the necessary background for simulation. In general, although robot teams are used, the configuration can be sufficiently described as a single robot problem, multiplying entities of identical type and capabilities.

Rules of the agent society (game)

In general the embedded agents are situated in a very simple, wide, but limited area, that is more abstract and contain less objects and forms than domestic living space. For this game like simulation, the control architecture must be indicated with an additional set of rules that go along with the game rules themselves. For the world knowledge, giving necessary constrains of the working environment, the semantic memory has contain rules based on the following assumptions:

- There are mandatory physical rules.

  Consequences: mechanical stress, temperature imbalances, and other influences on the robotic body causing bodily needs, which are basis for drives in pre-decision, can be simulated.

- To stay "alive", agents have to find energy resources.

  Consequences: this is a major impulse to act for the agent and helps to cause observable behavior.

- The total energy (environment and agents) is limited.

  Consequences: this allows the determination of a predictable end for the simulation experiment.

- In order to stay "alive", agents have to find energy resources and "feed" on them.

  Consequences: this invokes, that a agent has to change positions, and explore the environment.

- An agent always consumes energy. The level of consumption increases with its level of activity. If the agent's energy falls below a certain threshold defining the minimum energy level, it "dies".

  Consequences: in combination with the assumption before, this leads to an end of the simulation, and also allows, that the number of active agents is changing during the simulation, and

imbalances of agent teams can be simulated (e.g. a single agent competes with a group of agents)

− The simulation does not end until each member of one group has died.

Consequences: this allows a qualitative result, defining a winner and a looser group. Winner does not mean that all members have to survive. The goal of the group is that as many team members as possible, but at least one agent of the team can survive longer than the all agents completive group.

− Optional: collected objects do not inherit energy, that can be retrieved.

Consequences: this allows that the objects can be differentiated from energy sources in their maintenance and reward, emphasizing conflicts. An extra rewarding system can appropriate, but is not essential.

− A agent can help team mates:

Consequences: so it is useful for agents to cooperate.

Two groups of agents act competitively, trying to gain advantage over the energy resources. The goal of each agent is to live as long as possible and to cooperate with the other agents of its group to keep them alive, too. Furthermore there is a set of rules used in the memory system "Superego", defining cooperation abilities within the team.

− The agent has to survive as long as possible.

Consequences: internal imbalances or potential threats should be avoided.

− The sustainability of team members is as important as the own sustainability.

Consequences: an agent has to help team mates and the team agent can expect help, in case of asking for help.

− The help of competing agents is not granted.

Consequences: an agent cannot expect help from the other agent group.

The individual agents can differ slightly from each other as specified by different values of emotional parameters, different memory entries, or different available action patterns. This is a matter of configuration of the control system itself, as it will be described in Chapter 2.5. Thus, various configurations of the model are evaluated and compared to each other in search for an optimum behavior.

7.4 Environment configuration

For further testing, an evaluation of the major factors and settings of the agent and its environment simulation have to be determined. As proposed above, the agent has to show its ability to complete a mission in a game-like situation. Therefore, a minimum set of functionalities of the agent is required among the range of possible actions within the abstract environment. This requirement is directly linked with the characteristics of the "world," containing assumptions about basic properties, e.g. gravity, temperature, topography, morphology of mobile and immobile objects. The following settings are used for the simulation environment in further testing:

− Universal physical laws especially valid for the situation given on earth, e.g. acceleration of gravity, thermodynamics and temperature ranges, conservation of energy, mechanical deformation, mechanics, etc., are assumed in this environment.

Consequences: all movements in this world are supposed to "consume" energy. Heavy objects cannot be relocated without effort, robots are earthbound and movements can be viewed two dimensional positioning.

Affecting: movements and their effort are predictable, which is necessary to create bodily imbalances for drives (Chapter 3.1.4)

− The world is illuminated allowing visual perception of objects. Colors give minor information.

Consequences: visual sensors can be used and are highly desirable.

Affecting: this information is necessary for perception and definition of abstract images and their characteristics (Chapter 3.1.1)

− The world is not endless, but wide compared to the size of agents.

Consequences: the probability is high, that objects exist at least in far distance. Therefore the definition of an action radius smaller than the entire world appears necessary. Furthermore it inherits the consequence, that the agent cannot perceive all processes and objects in the environment.

Affecting: this definition effects directly the action planning. E.g. the seeking system has to indicate and localize necessary objects in the world by wandering and foraging. Cooperation and sharing knowledge within group can be emphasized.

− The world is highly static and silent. Only the agents themselves are able to move.

Consequences: this entails, that acoustic sensors can deliver only marginal information and might be not considered. Furthermore an agent can expect that objects out of sight are still on the same place, except when moved by other agents.

Affecting: this has great importance for memorizing and recognition of objects (Chapter 6.6.1)

− The world is flat without extreme inclination and unevenness.

Consequences: this allows a better overview of the local situation with a lower probability of hidden objects.

Affecting: perception becomes less uncertain. It is assumed that everything within the action ratio of the agent is perceivable (except when covered by another object). Furthermore the motion primitive has no restriction, as the agent can move to any position in the world with the same effort.

− Temperatures are in the range of -10 – 40 °C. The temperatures can change and can be unsteady in distribution.

Consequences: there are areas that are preferable to others. There might be areas on the playground that are harder to explore than others.

Affecting: this has a direct affect on drive vector (analyzed in Chapter 3.1.4 and defined for this approach in Chapter 6.2.4), which uses temperature imbalances. An agent of a high temperature imbalance will be not capable to explore this area at all. Beside drives, there are also emotions of the emotion vector (whose functionality is discussed in Chapter 3.1.5), as the seeking system that is influenced by these drives and competed with the other three emotional systems.

– The amount of energy in the world is limited.

Consequences: as soon as all utilizable energy is "consumed" (transformed into heat), the entire world stops (freezes).

Affecting: this is also determined in the drive vector, (analyzed in Chapter 3.1.4 and defined for this approach in Chapter 6.2.4), which indicates the energy resources of the robotic body. This has indirect effect on emotion systems.

– Energy resources are physically represented by container-like objects. There are different types of containers, which give access to a single agent or to two or more agents at the same time. The distribution of these resources is arbitrary.

Consequences: as energy is not arbitrarily accessible, the agent has to look for energy and has to manage resources in order to stay operable.

Affecting: the emotional seeking system (Chapter 3.1.5) indicates the exploration the world creating a cartography of potential object, that shall help to estimate the general situation of the environment (rich of energy resources, preferable temperature ranges etc.) for further assumptions.

In general, most properties closely resemble earth-like living spaces, as it is expected that the agents equipped with this software architecture are used in human living spaces. However the abstract level allows a faster decision making process.
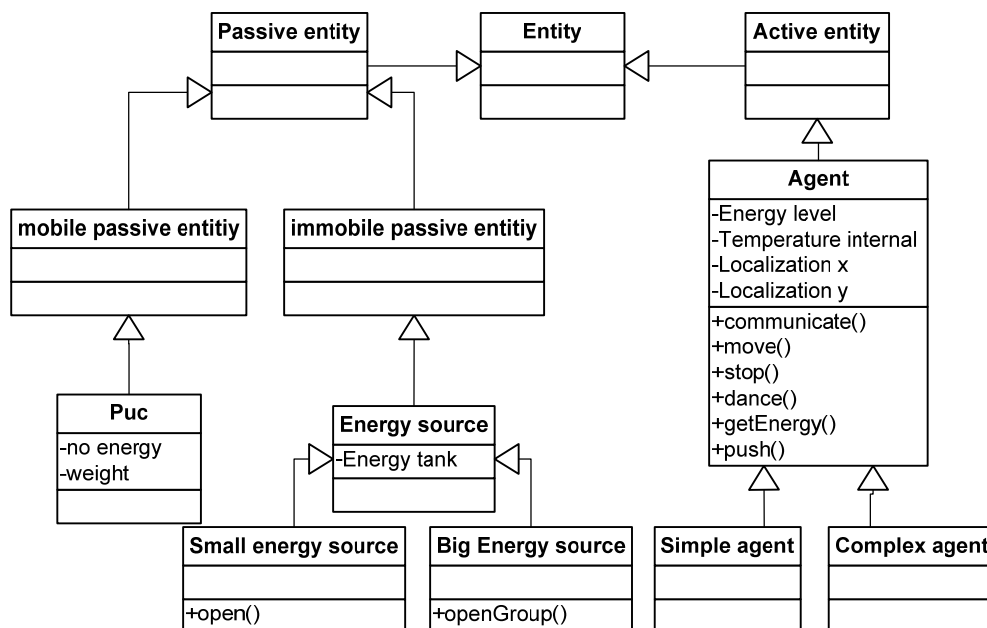

Figure 7.3: Entity classes of environment configuration

Figure 7.3 shows a general overview and dependencies of entities which can exist in the current environment configuration. In general there are two types of entities: passive entities cannot move, ore

more precisely change their position themselves (mobile passive entities) or at all (immobile passive entities), passive entities have no actuators, and in general also no sensors, as they represents the objects of this world. The objects of the world possess different attributes, e.g. energy sources inherit a certain amount of energy and provide operations to retrieve the energy by one or more agents, while other objects like "puc" which does not contain any energy to be retrieved and can be either seen as an obstacle or be used for foraging depending on eh rules of the game (depending on the scenario that shall be set up for testing). The second group are active entities, which in the current configuration are all agents, all agents posses the same physical properties and set of actions, but their control architecture can be either very simple and rule based (containing solely the reflexive control system of the model) and far more complex, inheriting the implementation of the complete model (complex agent). In simulation two homogenous groups of these two types of agents compete for the resources in this environment. The idea, that active entities like agents of different size and abilities can be seen as energy resources to emphasize conflicts between teams allow new aspects like persecution and hunting. Although this is not part of the current implementation, the shown class model of Figure 7.3 can be enhanced in this context.

7.5 Robot configuration

The general appearance and physics of the robot, simulated as an embodied agent, has been left out in this research in order to broaden the range of applicable robots. Under these circumstances, this aspect gains importance, leading to a paradox: the identical model in a different mobile robot might gain different results in the evaluation of its humanoid property. A robot of humanoid appearance, even if armed with a very simple behavioral schema, might be more readily accepted than an "intelligent" cubicle robot. The dilemma lies in the fact that external observers always estimate the overall robotic system and not just the model itself. The best solution still seems to be using a third party observation by using software agents rather than real robots so as to have a higher focus on the model itself without dependencies on physical properties and to achieve a higher level of repeatability. To have competition under fair conditions the robotic teams have to be facilitated with the same robotic body, differing only in the control system, as it is shown in Figure 7.3: the robots used in this simulation are active entities of the class "agent". The shape and physical capabilities of the robot are kept relatively abstract, emphasizing the simple modular actions as used in the examples of Chapter 6.7 to allow a wider range of potential robots in the application. For the simulation, the agents are equipped with a virtual set of sensors predefining the potential environmental information. Shaping the robots by using the small robot Tinyphoon[19] as an archetype, a simple set of robotic actions is allowed increasing the probability of evaluations in real world using this robot. Therefore the following properties have to be considered:

- − Sensors:
    - o Environmental sensors: visual, gravity, temperature and acceleration are a direct results of the assumptions of the environment configuration (Chapter 7.4), that shall be

---

[19] The robot is designed by a research of the same institute like ARS project. The robot has been originally designed for FIFA robot soccer and is comparatively simple in its form and actions. Further description of this project can be found in chapter 5.4.4.

kept close to real world situations. As there are light, gravity and temperature fields supposed to be detectable in the simulated world, the sensing of these types of information appears highly desirable.

- o Internal sensors: Beside external perception, internal values can be used for internal values indicating the states of the robotic body: Energy level has to be known in order to guarantee, that no lack of energy leads into operational disability. The inter temperature and the magnitude of deviation to the optimum is important to emphasize drives which lead to decision that might change unwanted internal states. Position and torsion of physics allow the estimation of potential damage, which should be prevented at all costs.

- − Communication: tools for simple, low-rate communication by bit-by-bit transfer. Communication protocol requires a sender and a receiver address. This is necessary to get in contact with team mates and provide the basics of communication for cooperation (asking for help). However, as it is not expected to transfer high amounts of data, e.g. to share new abstract images (created through experience), and no real time communication seems to be required (Only agents close to the broadcasting agent are supposed to get involved, and there is no time critical information exchanged considered in this state of simulation).

- − Mobility: ability to move on flat land with a medium inclination. Although the ability to change localization is seen as inevitable for the embodiment of the controlled agent, the method for movement is arbitrary, and can be walking, crawling, rolling, sliding etc. (the form affects only the lower motion primitive and the amount of consumed energy in execution)

- − Actuators: The amount and complexity of actuators is arbitrary. In general is the ability to move the only mandatory action in the first place. In case of applying further, more specialized tasks, e.g. transporting of special objects, maintaining or any form of handling objects might require further actuators and requirements on the physics of the robot. An example in this context has been given for service robots in domestic use in Chapter 2.3.

External sensors and actuators are the interface between the embedded agent and the environment, it is situated in. For the model the internal sensing is crucial, to give evidence about the "health" of the robotic body. While internal sensing allows assumptions about the balance of the internal systems and their needs that result into drives on more abstract level, the external sensing provides information about how to satisfy these needs in order to retrieve a balanced optimum. In analogy to the rather simple physical shape of the robot Tinyphoon, that has no apparatus to lift and carry an object, the basic actions of the example given in Chapter 6.7 has to be reduced to "move" and "push" an object.

The main difference between agents is not emphasized in their embodiment, which has to be unified, but in their control. One group of agents can make use of the full model (described in Chapter 6.2 and 7.6), there are more simplified rule based control architectures, repeating stereo typed behavior patterns, as they are presented exemplarily in Figure 7.4.
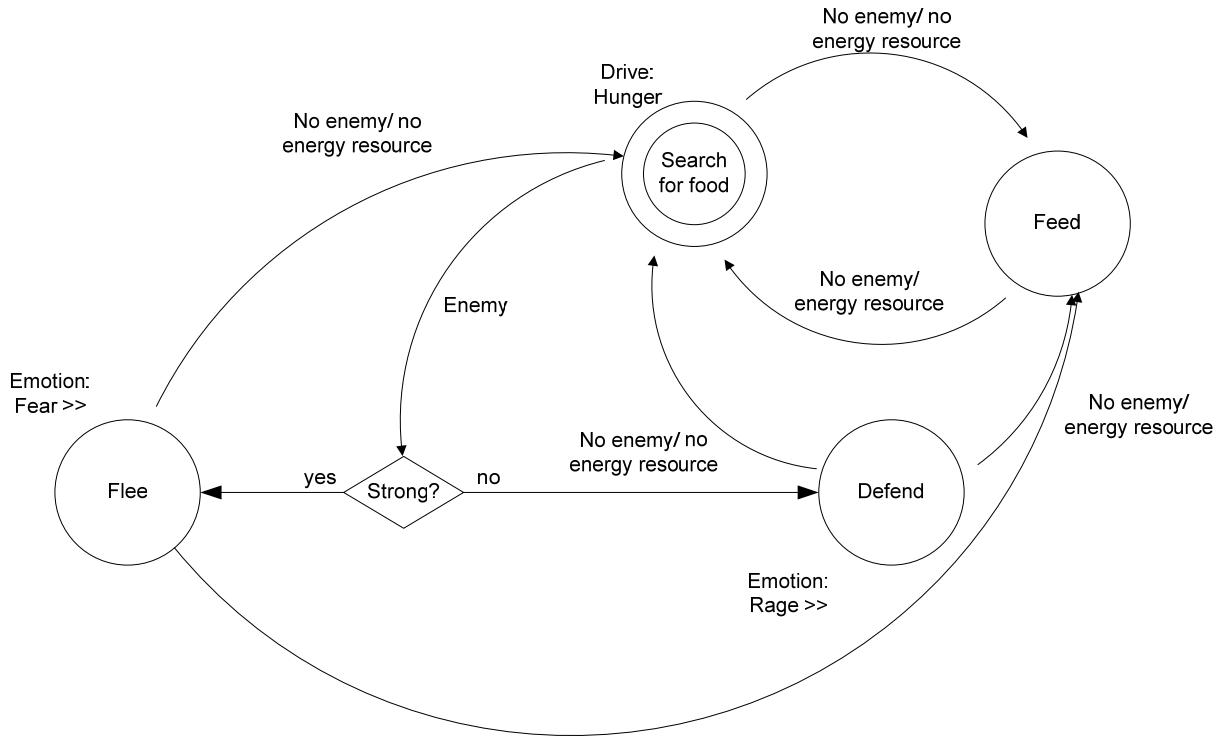
Figure 7.4: State chart of an exemplary stereo typed behaviour used for simple control

7.6 Model configuration

Besides an extensive configuration of semantic and application dependent rules for the physics of the model itself has to be configured, as it has been discussed defining this approach (Chapter 2.3). To optimize configuration number and type of values used of different functional entities have to be discussed. The functional model allows the change of diverse factors, e.g. drives, emotion vector, which can be used for configuration achieving optimization for control. The following chapter gives a brief introduction to the general values in design, which have to be kept in mind considering necessary constraints and implications. While pre-defining a set of images and episodes, which represent the experience, which shall be emphasized with learning methods in future, the settings of the drives and emotions are directly linked with the functional concept of environment evaluation (primary emotions), cooperative tendencies (secondary emotions) and the simulated robotic physics (drives). The episodes, which are either a part of experience and environmental knowledge, but also have a high impact on the functionality of modules directly, will be described separately in the test cases.

7.6.1 Abstract images (templates)

Abstract images (*Definition 3.3*) are the basic entities of the memory system, which are compared with the perceived environmental information compressed to an perceptual image (*Definition 3.2*). This basic set of these templates are always application-dependent. They are the configuration of the system, similar to the innate functionalities of an organism. They have to be adapted:

- Missions: defining the main tasks of objects and events expected for this mission. These are just simple crude motor behavior patterns, similar to reflexes.

Table 7.1: Exemplary set of images

Image (for cooparation test case)	(Primary) Emotional Vector					Drives			Bubbles visible	Energy source
	Seek	Panic	Rage	Fear	Lust	Hunger	Thermal imbalance	Stress		
Become Hungry					calculated					
no mate, e-source and no enemy	Low	Low	VeryLow	Medium		High/VeryHigh	Not set	Not set	0	true
no mate, no e-source and no enemy	VerHigh	Medium	Medium	High		Low/VeryLow	Not set	Not set	0	false
no mate, e-source and enemy	Medium	VeryHigh	VeryHigh	VeryHigh		Low/VeryLow	Not set	Not set	1+	true
no mate, no e-source and enemy	VeryHigh	VeryHigh	Low	VeryHigh		Low/VeryLow	Not set	Not set	1+	false
mate hungry, e-source and no enemy	VeryLow	Low	High	High		High/VeryHigh	Not set	Not set	1+	true
mate hungry, no e-source and no enemy	VeryHigh	High	Medium	VeryHigh		High/VeryHigh	Not set	Not set	1+	false
mate hungry, e-source and enemy	High	High	Medium	High		High/VeryHigh	Not set	Not set	1+	true
mate hungry, no e-source and enemy	Low	VeryHigh	VeryHigh	VeryHigh		High/VeryHigh	Not set	Not set	1+	false
mate not hungry, e-source and no enemy	VerHigh	VeryLow	High	High		Low/VeryLow	Not set	Not set	1+	true
mate not hungry, no e-source and no enemy	Low	High	Medium	High		Low/VeryLow	Not set	Not set	1+	false
mate not hungry, e-source and enemy	VeryHigh	VeryHigh	VeryHigh	VeryHigh		Low/VeryLow	Not set	Not set	2+	true
mate not hungry, no e-source and enemy	Medium	Medium	VeryHigh	High		Low/VeryLow	Not set	Not set	2+	false
mate not hungry, no e-source and enemy	VeryHigh	High	VerLow	High		High/VeryHigh	Not set	Not set	2+	false
Not Hungry										
no mate, e-source and no enemy	Medium	Low	VeryLow	VeryLow		Low/VeryLow	Not set	Not set	0	true
no mate, no e-source and no enemy	High	Low	VeryLow	Low		Low/VeryLow	Not set	Not set	0	false
no mate, e-source and enemy	Medium	High	High	Medium		Low/VeryLow	Not set	Not set	1+	true
no mate, no e-source and enemy	Medium	High	Medium	Low		Low/VeryLow	Not set	Not set	1+	false
mate hungry, e-source and no enemy	High	Low	Low	High		Low/VeryLow	Not set	Not set	1+	true
mate hungry, no e-source and no enemy	Medium	High	Low	Medium		Low/VeryLow	Not set	Not set	1+	false
mate hungry, e-source and enemy	Medium	Medium	VeryLow	VeryLow		Low/VeryLow	Not set	Not set	1+	true
mate hungry, no e-source and enemy	Medium	Low	VeryLow	Low		Low/VeryLow	Not set	Not set	1+	false
mate not hungry, e-source and no enemy	Low	VeryLow	Low	Low		Low/VeryLow	Not set	Not set	2+	true
mate not hungry, no e-source and no enemy	Low	High	VeryHigh	High		Low/VeryLow	Not set	Not set	2+	false
mate hungry, e-source and enemy	Medium	High	High	High		High/VeryHigh	Not set	Not set	2+	true
mate not hungry, no e-source and enemy	Medium	Low	Low	Low		High/VeryHigh	Not set	Not set	2+	false
mate not hungry, e-source and enemy	High	High	Medium	Medium		High/VeryHigh	Not set	Not set	2+	true
mate not hungry, no e-source and enemy	Medium	Medium	Medium	Medium		High/VeryHigh	Not set	Not set	2+	false

- Proposed environment: showing special application depending properties e.g. domestic use, factory site, open air, contact with humans, contact with animals, contact with other machines.

- Used robot: defines sensing information based on which sensors are available, mobility (top speed), and internal values considered.

The purpose of the used robot specifies the set of actions: a rescue robot *must not* run away, a cleaning robot cannot be used for lawn mowing, etc. Images are directly connected desire-plans, which transitions are indicated by the images. Desire-plans themselves also result from the same three assumptions, as they build a link between the environment in cooperation with robot dependent needs and tasks. In general linked to the images, which are indicators of state transitions in multiple desire plans, their indication and exact description are therefore crucial for the success or loss of performing a task invoked by desire plans. For the test cases described in the following chapters a sample set of states have been built and inserted into the image memory (described in Chapter 6.6.1) These images will not be changed during simulation, as they appear efficient for this test case and guarantee higher predictability in perception. The example set shows how the emotion vector (represents the four basic emotions, which have been emphasized in Chapter 6.2.4 described in the following chapters.

Although drives are not part of the environmental perception and therefore not part of the association process described in Chapter 6.2.3, it is decided for simulation to integrate the drive vector in images in order to overcome the problems of missing embodiment in simulation, which might lead again to a too hardware dependent solution interpreting robotics physics too tightly.

7.6.2 Emotion vector

As described in Chapter 3.1 emotions and drives represent very important evaluation systems of the human mind. However their categorization and interplay do not seem to have a unified interpretation in psychoanalysis and cognitive science, which has been shown in Chapter 5.3.1, summarizing the main theories in this context. However there is high evidence that the theory of J. Panksepp (Chapter 3.1.5 and 5.3.1), which has be seem to be very appropriate, proposing 4 primary emotions, and will be used for this approach (as described in Chapter 6.2.4): the emotion vector is split into four primary emotions and seven secondary emotions. The four primary emotions seek (curiosity), fear, rage and panic refer to Panksepp's theory (Chapter 3.1.5). They are the basis and directly linked to a selection of complex (secondary) emotions, which represent an appropriate set for situation-dependent evaluation. There are seven different complex emotions referring to four emotional control systems (Chapter 3.1.5)

- *Hope*: indicates the possibilities of successfully finishing the desire plan.

- *Joy*: shows that a good event has happened.

- *Disappointmen*t: represents that a bad event has happened.

- *Gratitude*: increases when another agent (of the team) has done something good to the agent.

- *Reproach*: arises when another agent (of the team) has done something bad to the agent.

- *Pride*: results as soon as the agent itself has done something good (to another team agent).

- *Shame*: is increase when the agent has done something bad (to another team agent).

Hope is the only a-priori emotion, which rises in expectation of an event. It is an indicator for the probability of the occurrence of the expected events. The other six secondary emotions give an evaluation of past events. This set of secondary emotions shall give evidence to define the social behavior of the embedded agents in team, and indirectly go along with the rule set of the Superego, which as been defined in chapter 7.3. The main difference to primary emotions is not only that secondary emotions are an interpretation of different times (past events or future outcome), but also their influence on robots in team. Using social respect as a direct value, which can be increased not by itself, but by *other* team members, allows to implement a form of ranking in the group. A high value is very desirable, as it increases the portability to receive support by team members, and indirectly the probability to survive. In general two forms of social respect can be implemented: the direct social respect ($r_{ij,}$= the respect of member *i* towards member *j*) between two members of a team, that can range between -1 and 1 (0 is neutral, and means equality) and general social respect (r_g), that causes the absolute ranking (*i*) within a group (*i* in Equation 7.1 and gives a positioning where 1=top position, n=least position). Direct respect can affect the cooperation of two concrete team members in the current situation, while the general social respect determines a group hierarchy in the group, and can be derived as follows:

$$r_g = \sum_{i=1}^{n} \frac{r_{ij}}{i} \qquad \text{for } -1 \leq r_{ij} \leq 1 \qquad (7.1)$$

7.6.3 Drives

The drive vector, as described in 7.6.3, shall be used in this simulation, and can be found in image storages of the test examples of this simulation. Three drives are part of the vector, while the fourth drives is calculated sum, giving the absolute value of the vector (described in Equation 6.8):

− Hunger: the level of hunger indicates the shortage of energy.

− Thermal imbalance: high temperatures indicate the malfunction of mechanical or electronic devices on the robot.

− Physical imbalance: weight sensors can indicate the balance of the robotic body indicating the stabilization of current motion primitives.

− Lust: a general indicator of the inner state of the robot.

Besides stored values, giving an interpretation of the supposed outcome of an episode or meaning of a single perceptual image (*Definition 3.2*), further information affecting the interpretation of perceptual image, internal values take direct influence on the amplitude of the vector. The calculus of the vector can be described as follows: the drive value "Hunger" (*H*) is indirectly proportional to the current energy level (*e*) of the robotic body. The drive can be triggered with following equations depending on the level of energy (*e*):

$$H = e_{ideal} - e / e_{ideal} \qquad \text{for } e \leq e_{ideal} \qquad (7.2)$$

$$H = e - e_{ideal} / e_{ideal} \qquad \text{for } e \geq e_{ideal} \qquad (7.3)$$

The drive thermal imbalance (I_T) is triggered by out-of-range sensory values, defined by a minimal and maximum temperature. The degree of variance is proportional to the strength of the drive. As the robot may consist of a number of critical parts, the drive is dependent on an array of temperature values.

$$I_T = \mid T - T_{\max} \mid \qquad\qquad \text{for } T_{\max} \leq T \qquad\qquad\qquad (7.4)$$

$$I_T = \mid T_{\min} - T \mid \qquad\qquad \text{for } T_{\min} \geq T \qquad\qquad\qquad (7.5)$$

Stress as physical imbalance cannot be set up as a general parameter, as it is directly coupled with the physics of the robotic body. In the case of the current simulation, it will not be considered at the moment. The lust drive (*L*)is a combination of the other three drives and acts inverse to the others (its calculation has been described in Equation 6.8 of Chapter 7.6.3). If the other drives are low, lust is supposed to be high and vice versa.

7.7 Testing in simulation environment

The following chapter shall emphasize the concrete settings and scenarios used for the testing of control based on the functional model. Therefore a sub set of scenarios shall be tested on two teams of agents. For testing the same physics of agent different control systems are applied to achieve observable differences in behavior. The following chapters describe three exemplary scenarios in order to categorize the behaviors emerged in these tests, which have been described first in [Deutsch 2007].
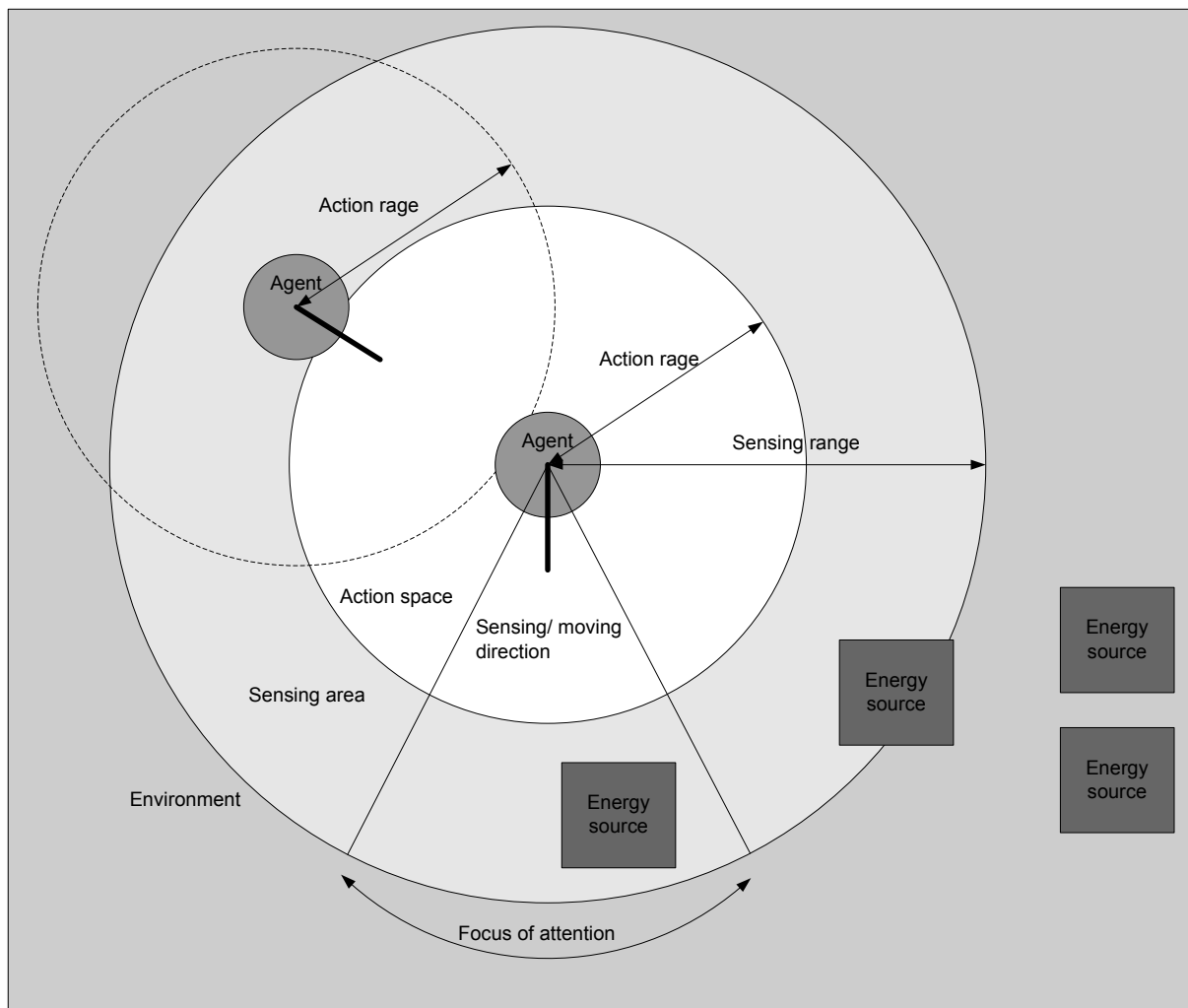
7.7.1 Defining scenarios

Figure 7.5: Sensing ratio and action space of agent

The main question in simulating the interplay of environment robot guided by the control system is how to define abstract and reproducible situations, building an abstract, reproducible scenario (defined Chapter 3.1.3). In this context the following proposal is made: as the embodied agent is supposed to possess actuators and sensors, I define a sensing radius and action space. The sensing space is supposed to be wider than the action space, and represent the agent's perceivable part of the environment. The size and form of the sensing space will depend on the sensors and their properties, important is that within this sensing ration there might be areas of higher interest and better perception depending on the (simulated) physics of the robot. All objects beyond this radius are "invisible", "nonexistent" for the agent, except their localization is memorized. The action radius depends on the actuators and motion tools of the embodied agent: entities (passive or active) within the action space can be sensedand can be affected by the agent "immediately". These objects or other agents are considered to have an impact on the current scenario as they allow direct interaction. Based on the basic assumptions, which are graphically shown in Figure 7.5, three basic scenarios "Ask for dance", "Cooperation for Food" and "Call for help" have been set up, providing the robot with all necessary configuration, going along with the evaluation test of this project represented by [Deutsch 2007]. The last scenario "Call for Help" can be a sub scenario of the "Cooperation for Food", showing the connectivity of different scenarios.

7.7.2 Ask for dance

This simplest form of scenario requires, a second agent (team mate), and is only founded in social rewards, gaining higher levels on secondary emotions for a higher ranking. To gain social respect within the group, it is desirable to come close to team mates and interact with them in a positive way. A agent, that is willing to play will broadcast messages, asking for dance and will try to attract team mates within the sensing ratio. Depending on its social ranking and direct social respect of the other team mate receiving the message dancing will be executed or not. This scenario shall emulate a contrast to the seeking systems, creating minor conflicts of behaviors of low priority (wandering, seeking) in non emergency cases. Not to dance with a team mate can lead to the lowering of social levels, and cause a change in secondary emotions (of both agents).

The desire for dance is randomly invoked, under the precondition of a high lust drive (which refers, that the current health of the agent is very balanced). The call for dance can be either start under the condition of sense a team mate or be activated arbitrarily.

7.7.3 Cooperation for food

In this scenario, two bubbles have to cooperate to open an energy source (big energy source of entity classes in and receive the required energy. In is a more complex scenario within this test sequence, assigning two separate roles: an agent, which needs help and a helping agent. There are two main roles in this scenario:

- A robot that seeks energy as its own stored amount is under a certain level: it has to ask for help, and has to wait for a helping team mate.

- An assisting robot giving access to the energy source: it has to decide either to help, or not and in case of help give feedback, join the team mate and execute the process.

Figure 7.6 and Figure 7.7 show the necessary episodes of this scenario. Using parallel states as described in Chapter 6, a robot can have these scenarios. There are competing scenarios, which cannot be initialized in parallel as they are based on different per-conditions describing general settings of the environment. Only the simplest episode consisting only of "init" and "hungry" does not require any pre condition and can be executed parallel, as it is based solely on inner functions of the system. The initialization depends on pre assumptions of the general situation, which can be statistical counters, e.g. only few team member exist, but there is plenty of energy in the environment, a state chart, proposing, that only in case that the robot comes no matter if close or not to the robot, as the probability to retrieve help is minor, and the risk to loose too much energy by helping can be decreased as the energy resources are sufficiently rich in the current setting. These setting are not pre defined, but can be advised to the agent by setting counters, giving sensing rates. E.g. in case a great number energy sources (exceeding a limit set by configuration) have been seen in the last time periods, the control system can assume that this environment is full of resources, and vice versa.
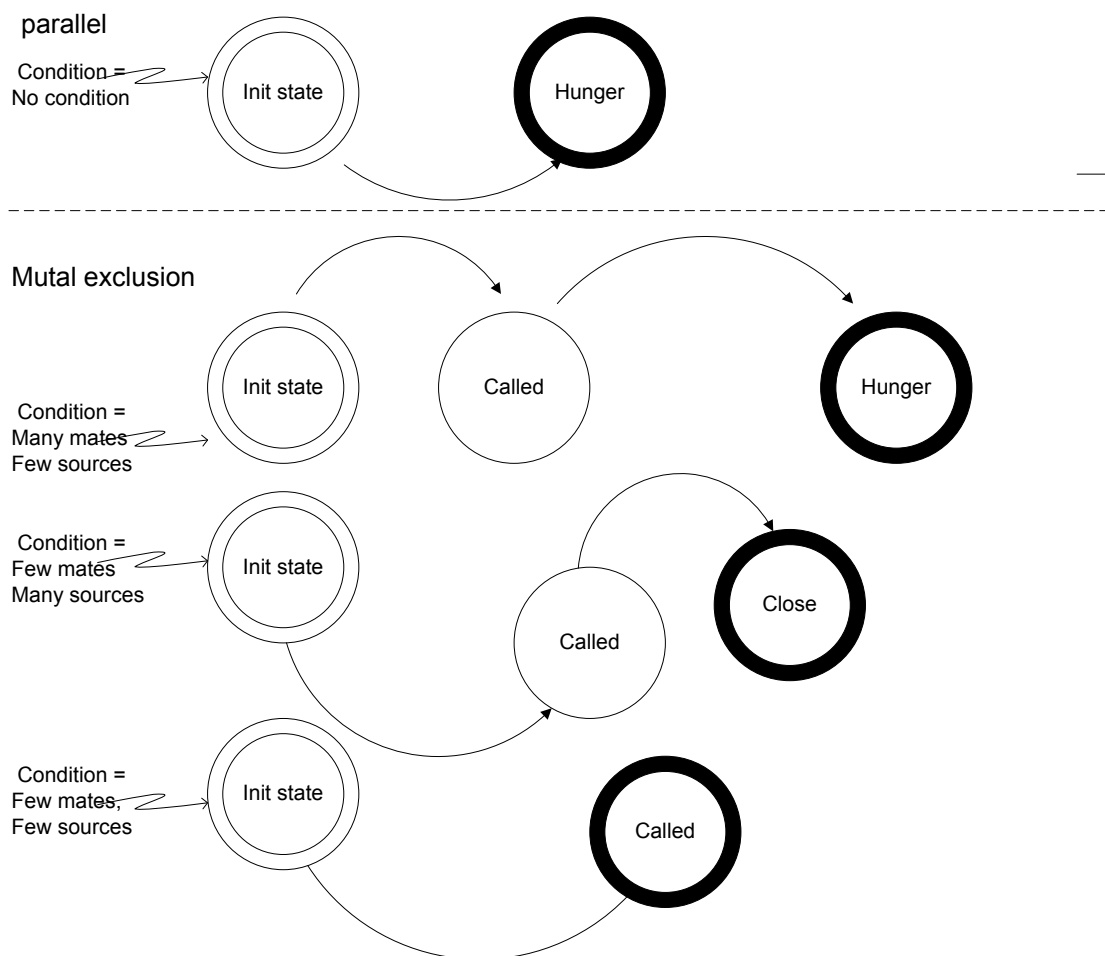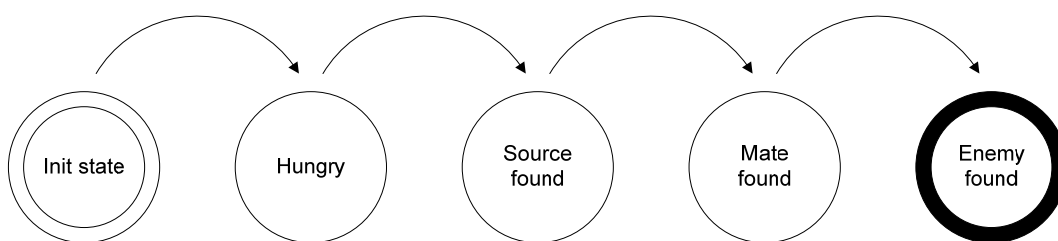

Figure 7.6: Episodes for test case 1 in working memory

Figure 7.7: Additional episode

7.7.4 Call for help

The scenario "Call for help" can be either a part of the scenario "cooperation for food" or start in other emergency scenarios. Otherwise than "Ask for dance" here the priority to help is a lot higher and a reward of higher rage. "Call for help" can be initialized, if one or more foreign agents are reaching the sensing area of an agent, or in case a big energy source has appeared in the sensing range, and the agent is required to open it in cooperation with a team mate. In latter case it is part of the scenario "cooperation for food". However the scenario "Call for help" can be detected separately.

7.7.5 Visualization

In general the system posses three different forms of visualization: 3D visualization, which is equal to a 2d visualization, showing the whole environment the robot teams are placed in, whit the ability to highlight the internal values of a agent selected by the user. Furthermore an extra test bench has been designed, testing the inner functional blocks.

The purpose of the test bench is to explicitly check the decision making of the control architecture, identifying episodes initialized in the working memory, the timing diagrams of emotion vectors and the selected actions. There are three situation defined within the BFG which shall be used for evaluation.
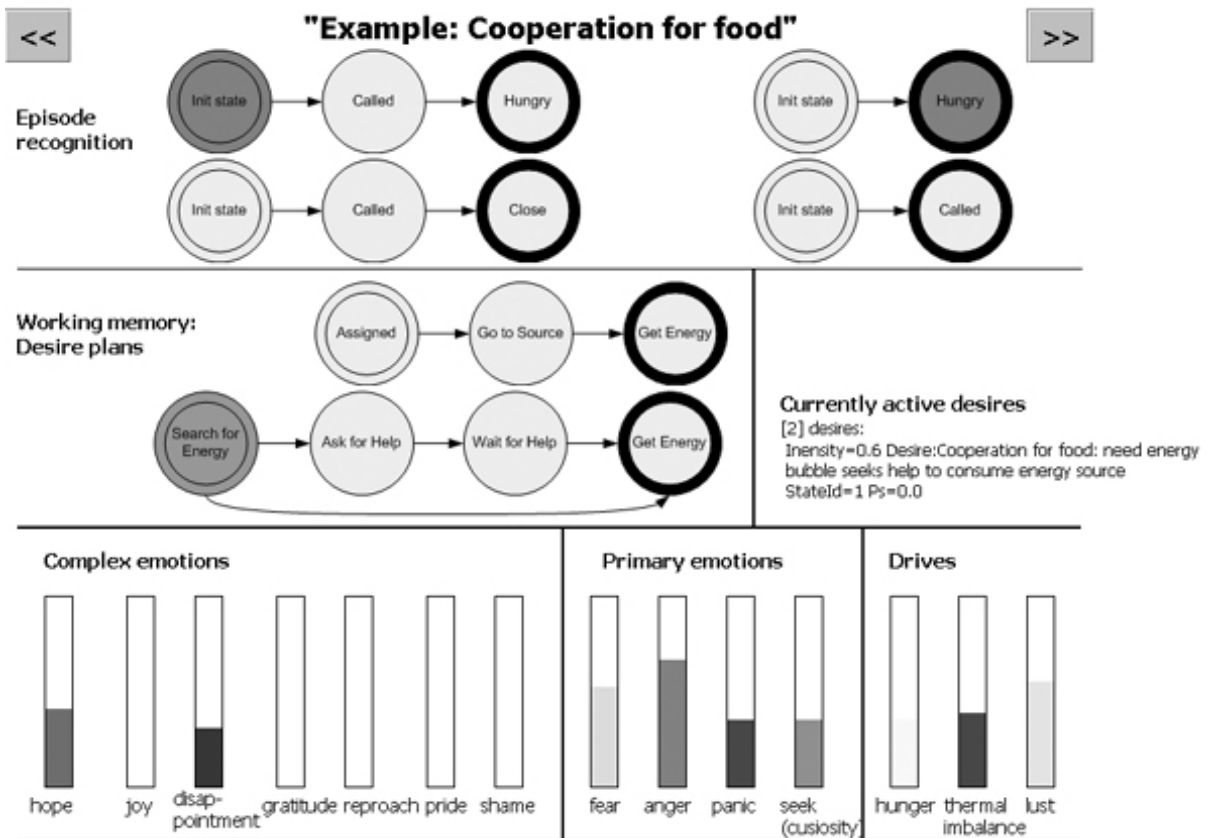

Figure 7.8: View box of test bench

7.7.6 Tests set ups and results

First, evaluative tests have been made using two competitive teams of four robots each. Each team contains of the same number of robots, with the same embodiment, to achieve the same abilities to move and act on the "play ground". The only difference between the teams lies in the mental abilities

of the robots. One of the teams uses the first implementation of the complete model, and the opponent team is facilitated with a simple rule-based control of the model as presented for the reactive control system of the model. This rule-based control system can show simple behavior schemes based on if/then clauses, providing similar stereo type behavior as shown exemplarily in Figure 7.4 in detail. These robots can show simple behaviors, e.g. "promenade" (containing of basic action "move" of Chapter 6.7, no observable scenario) composed by act on simple actions, as defined in the chapters 7.5 and 6.7.Furthermore this control is capable to take into account external influences and one internal parameter, the energy level.

The tests are based on three principal scenarios, presented in the chapters above, which have been designed for the game like designed test environment have been published first in [Deutsch 2007]. The following scenarios have been used for test:

− Ask for dance: this case it is activated, when a team mate is perceived and no internal imbalances (drive vector) exist.

− Cooperation for food: this is a more complex situation, where energy resources require cooperation for opening.

− Call for help: this is a general situation, where a team mate is asked to come as soon a foreign robot is recognized as a common threat.

Table 7.2: Survival rate after1000 cycles [Deutsch 2007]

	Simple agent	Complex agent
Min number. survived	0.00	1.00
Average number survived	2.79	3.44
Max number survived	4.00	4.00
Standard deviation	0.52	0.65

To shorten the test runs, the simulation was not terminated by death of the whole group, but operation cycles of 500, 1000 and 2000 for each sequence of (accelerated with approximately 1:1000) simulation runs. The survival rate (number of robots per team) for 1000 cycles is shown in Table 7.2. Although both teams may loose members, the complex agents can survive in greater number, and have never completely died out in simulations of 1000 cycles and 2000 cycles. In general the robot team with a complex behavior architecture shows a significantly longer survival time (about 60% longer than the simple robots) [Deutsch 2007]. Furthermore there is an interesting phenomenon, indicating that the probability of a complex agent to die decreases with simulation time, while simple agents have a constant probability. This is linked with the ability of adaptation.

The test bench, which shows the internal processing of the complex agents, gives the state and the used episodes. For this test the scenarios (Chapter 3.1.3) can be directly referred and displayed with an abstract episode (*Definition 3.7*) set for the simulations can be referenced to (abstract) episodes, which are visible in the test bench and can be used for further qualitative evaluation. The evaluation sequence with 25 runs of 1000 cycles is shown in Table 7.3. As it is possible that more than one scenario can be

activated at the same time due to the different states of agents in team, the sum of the scenarios does not necessarily be normalized to 100%. The test show that the episode No. 1 (indicating the possible scenarios for this test) "Ask for dance "is activated in 48,8%, and appears most popular although it has only social benefit. The simplicity of the scenario No.1 compared to episodes like No. 2 "Cooperation for food" might be a reason for this effect. The standard derivation of the scenario 2 lies in the potential loss of team mates, which are a necessary requirement in this scenario. In case team members died early in the simulation run, the "call for help" for this scenario remained without response and lead to Resets and break-off of this episode. In general it seems that social components provided by the higher functional entities of this model are in use, if provided. Although there is no direct benefit for survival, social ranking seem to have great influence on adaptability and seem to extend the probability to survive, as shown in the first part of testing.

Table 7.3: Scenarios activated in complex agents (1000 cycles)

Scenario:	No1: Ask for dance	No2: Cooperation for food	No 3: Call for help (and not part of No 2)
Min %in simulation time	48,8%	22,2%	0%
Average % in sim. time	76,1%	58,8%	6 %
Max % in sim. time	85%	95%	19%
Standard derivation	9%	21%	5%

8 Conclusion and outlook

The thesis describes a functional model for control systems of autonomous systems, with a special focus on service robots. The model contains novel concepts for decision making and action selection based on recognition and emotional evaluation of environmental information, retrieving goal-oriented behavior for autonomous task completion. The model is based on three different control systems allowing simpler and rule-based behavior of high performance in combination with complex goal-oriented behavior considering constraints and consequences of actions. The key features of this model are:

- Emotional evaluation and filtering in pre-decision module: allows the extraction of essential information
- Three overall control systems combined with a set of fixed rules
    - Reflective Control: can schedule action sequences based on memory for action planning and goal-oriented behavior
    - Routine Control: can be initiated by other controls and contains fixed (repeating) action sequences
    - Reactive Control: simple rule-based reaction appropriate for emergency cases and dynamics.
- Competitive emotional control systems: used for action planning based on emotional values, which represent the current internal state of the system.
- Conflict solution in action module: providing coherent behavior
- Memory systems and working memory: allow action planning and learning

The model was inspired by theories of psychoanalysis and cognitive sciences. In particular the emotional theories of [Panksepp 1998], [Solms 2002] and [Damasio 1999] had great influence on the design of this model. The main idea is to use a hierarchical architecture, which uses reactive control loops as well as reflective ones. The latter are inspired by an investigation into the higher functions of the human brain and take into account recent results in sciences that deal with the mind. Apart from sensors and actuators, drives and emotions complete the embodied character of this model. Providing the autonomous system or robot with drives activates the necessary explorative and motivated behavior in the first place. Emotions act as evaluation mechanism, coordinating inner and outer needs. There is special interest in the resolution of situations with contradictory behavioral possibilities. Social contexts requiring individual as well as team-oriented behavior are a good source for such situations. In biology, the need to act socially is regarded as an important factor in the evolution of intelligence. A crucial factor is the scheduling of partly competitive actions proposed by different control systems. This inner conflict can be maintained as a separate functional module in order to ensure consistent, coherent behavioral sequences. Using different long-term memory systems combined with a working memory for currently used data allow effective action planning and task decomposition. The simulation environment shall provide a qualitative analysis of performance and correctness of model. Differing from simple test environment, the environment in human habitation is far more complex and statistics cannot provide a sufficient evaluation schema that result in generally valid conclusions. The modular blocks of the simulation allow the adaptation to different application fields and the test of a

variety of different versions of control architectures, facilitating "evolutionary" progress in the design of autonomous behavior. By means of a special simulation module, the test bench, internal communication and state of the different functional blocks can be tested, delivering insight in the mechanisms and their mode of operation. This can give information about the observed behavior of the simulated agents using this model for decision making.The benefit of the functional model shall be qualitatively shown in situations, where competitive groups of robots using different control system try to gain advantage over the opposing team. A model of this complexity cannot prove its efficiency by exclusively quantitative analysis. Results in this context are hard to measure in quantitative terms, but using competitive situations makes it possible to show qualitative differences. Therefore, a game-like situation can show the advantages and disadvantages of a model in a given situation. The focus of the simulation lies in the provocation of conflicts and their potential implications, which are a challenge for other, normal behavior models. In predefined test cases providing a sequence of different scenarios, which are supposed to cause general conflicts, it can be tested which strategies are successful and how they can be initiated. The proposed model is the result of this evaluation process, as proposed in Chapter 2.

Testing intelligence is being discussed in different research communities. As [Bryson 2007] stated, the lack of new approaches providing behavior architectures for general purposes may be based on the fact that it proves to challenging to prove their correctness in evaluation. Especially in the case of emulating other forms of intelligence besides cognitive intelligence, which can be expressed more easily by mathematical abilities and therefore is more feasible in statistical measurements, the testability is controversial. Testing cognitive abilities is primarily aimed at performances analyzing the ability of problem solving. The basic idea is that the knowledge of cognitive abilities implies the ability to perform cognitive abilities. But this assumption is not necessarily adaptable to other forms of intelligence. An expert knowledge about emotions does not necessarily entail a skilled emotional behavior [Buller 2005]. This fundamental distinction is a main dilemma in psychoanalysis, which can be overcome by behavior observation rather than testing in case a number of pre-conditions can be fulfilled. The main challenge is correlation of personality and intelligence questioning how interpersonal emotional intelligence can be tested separately. Another obstacle is the emotional effects of the environment that influence the way of living. Most tests are situated in a life space-related dimension, compassing a tightrope walk between abstraction and completeness. This problem is even reinforced in the case of evaluating interdisciplinary research results, which have to meet the demands of two different approaches based on different fields of science, each of them insisting on divergent requirements in proving the correctness of the research. A robot is not capable of giving a self-reflective view in questionaires, nor can a complex model like this providing non-deterministic behavior be evaluated as based on quantitative criteria. The proposed evaluation method seems currently the "best practice" of proving in an abstract test environment as proposed by M. Toda [Toda 1982] the practical benefits of the model, similar to the habit in behavior research emphasized by J. Panksepp [Panksepp 1998].

The application fields of this model are autonomous systems with a special focus on autonomous service robots for the domestic area. In the future projects SENSE and SEAL (Chapter 5.4.4), the system shall be part of the control systems and mobile robots that will be applied to different services for in-field testing to prove the benefits of this model.

Bibliography

Prints media

[Aamodt 1994]   A. Aamodt, E. Plaza: "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", *AI Communications*, Vol. 7 (1), p. 39-59, 1994.

[Albus 1989]   J. S. Albus, R. Lumia, J. Fiala, Albert Wavering: "NASREM, The NASA/NBS Standrard Reference Model for Telerobot Control System", *Proceedings of the 20th International Symposium on Industrial Robots*, 1989.

[Anderson 2002]   J.A.D.W. Anderson: "Robot Free Will", *Proceedings of the 15th European-Conference on Artificial Intelligence*, ECAI 2002, IOS Press, p. 559-563, 2002.

[Arbib 2004]   M.A. Arbib, J.-M. Fellous, "Emotions: from brain to robot", *TRENDS in Cognitive Sciences*, vol. 8, p. 554-561, 2004.

[Arkin 1998]   R. C. Arkin: *Behavior-based Robotics*, The MIT Press, 1998.

[Baars 1996]   B.J. Baars, K. McGoven: "Cognitive views of consciousness: What are the facts? How can we explain them?", *The Science of Consciousness: Psychological, Neuropsychological, and Clinical Reviews*, M. Velmas (Ed.), Routledge, 1996.

[Baddeley 1997]   A. Baddeley: *Human Memory: Theory and Practice*, Psychology Press, 1997.

[Balch 2000]   T. Balch: *TeamBots, Software documentation*, Georgia Tech Research Corporation (GTRC) and Carnegie Mellon University (CMU), 2000.

                  Available (06/2007): http://www.teambots.org

[Bekey 2005]   G.A. Bekey: *Autonomous Robots – From Biological Inspiration to Implementation and Control*, The MIT Press, 2005.

[Blank 2004]   D. Blank, D. Kumar, L. Meeden, H. Yanco: " Pyro: A Python-based Versatile Programming Environment for Teaching Robotics", *Journal of Educational Resources in Computing (JERIC)*, No. 4, 2004.

[Bonasso 1997]   R.P. Bonasso, D. Kortenkamp, D.P. Miller, M. Slack: "Experiences with an architecture for intelligent reactive agents", *Internal Report Metrica Robotics and Automation Group*, NASA Johnson Space Center, 1997.

[Bolmsjö 1995]   G. Bolmsjö, H. Neveryd, H. Eftring: "Robotics in Rehabilitation", *IEEE Transactions on Rehabilitation Engineering*, Vol. 3, No. 1, 1995.

[Brainin 2004]   E. Brainin, D. Dietrich, W. Kastner, P. Palensky, C. Roesener: "Neuro-bionic Architecture of Automation Systems – Obstacles and Challenges", *IEEE Africon,* p. 1219 – 1222, 2004.

[Bresciani 2001]   P.Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, J. Mylopouslos: "A Knowledge Level Software Engineering Methododology for Agent Oriented Programming", *Autonomous Agents*, p. 648-655, 2001.

[Brody 2004]   N. Brody: "What Cognitive Intelligence Is and What Emotional Intelligence Is Not", *Psychological Inquiry*, 15, p. 234-238, 2004.

[Brooks 1986]   R.A. Brooks: "A robust layered control system for a mobile robot", *IEEE Journal of Robotics and Automation*, 2, p.4-23, 1986.

[Brooks 1991]   R.A. Brooks: *Intelligence without reason*, A.I. Memo 1293, MIT, 1991.

[Brooks 1998]   R.A. Brooks, C. Breazeal (Ferrell), R. Irie, C. C. Kemp, M. Marjanović, B. Scassellati, M. M. Williamson: "Alternative essences of intelligence", Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, p. 961-968, 1998.

[Bruckner 2007]   D. Bruckner, *Probabilistic Models in Building Automation: Recognizing Scenarios with Statistical Methods*, PhD Thesis, Vienna University of Technology, 2006.

[Bryson 2000]   J. Bryson: "Cross-paradigm analysis of autonomous agent architecture" *Journal of Experimental and Theoretical Artificial Intelligence*, p.165-190, 2000.

[Bryson 2007]   J.L.Bryson: "Mechanisms of Action Selection: Introduction to the Special Issue", *Adaptive Behavior*, Vol 15, No. 1, Sarge Publications, p. 5-8, 2007.

[Buffet 2001]   O. Buffet, A. Dutech, F. Charpillet: "Incremental Reinforcement Learning for Designing Multi-Agent Systems", *Proceedings of the Fifth International Conference on Autonomous Agents (Agents'01)*, 2001.

[Buller 2002]   A. Buller: Volitron: "On a Psychodynamic Robot and Its Four Realities", *Proceedings Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems 94*, p. 17-20, 2002.

[Buller 2005]   A. Buller: "Building Brains for Robots: A Psychodynamic Approach", *Invited talk on the First International Conference on Pattern Recognition and Machine Intelligence (PReMIT'05)*, 2005.

[Burgstaller 2004]   W. Burgstaller, T. Rausch: "Gatewayless Communication between LonWorks and BACnet", *Proceedings of the WFCS 2004 IEEE International Workshop on Factory Communication Systems*, p.361 – 364, 2004.

[Chapman 1985]   D. Chapman: *Planning For Conjunctive Goals*, Technical Report AITR-802, MIT Press, 1985.

[Cote 2006]   C. Cote, Y. Brosseau, D. Letourneau, F. Michaud, and C. Raievsky: "Robotic software integration using marie," *International Journal on Advanced Robotics Systems*, vol. 3(1), p. 55–60, 2006.

[Damasio 1994]   A. Damasio: *Descartes'error: emotion, reason, and the human brain, Penguine Books*, first published 1994, edition 2005

[Damasio 1999]   A. Damasio: *The Feeling of What Happens: Body amd Emotion in the Making of Consciousness*, A Harvest Book / Harcourt, Inc., 2003.

[Damasio 2003]   A. Damasio: *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, A Harvest Book / Harcourt, Inc., 2003.

[Deutsch 2006]   T. Deutsch, R. Lang, G. Pratl, E. Brainin, S. Teicher: "Applying Psychoanalytic and Neuro-Scientific Models to Automation.", *Proceedings of the 2nd IET International Conference on Intelligent Environments*, Vol. 1, p. 111 – 118, 2006.

[Deutsch 2007]   T. Deutsch, H. Zeillinger, R. Lang: "Simulation Results for the ARS-PA Model", *5th International IEEE Conference on Industrial Informatics*, INDIN´07, p.6, 2007, to be published.

[Dietrich 2000]   D. Dietrich: "Evolution potentials for fieldbus systems", *3rd IEEE International Workshop on Factory Communication System*, p. 145 – 146, 2000.

[Dietrich 2004a]   D. Dietrich, W. Kastner, H. Schweinzer, H.: „Wahrnehmungsbewusstsein in der Automation – ein neuer bionischer Dekansatz"; *at*, 52. Jahrgang, pp. 107-116, 2004.

[Dietrich 2004b]   D. Dietrich, W. Kastner, T. Maly, C. Roesener, G. Russ, H. Schweinzer: "Situation Modeling", *5th IEEE International Workshop on Factory Communication System*, p. 93 – 102, 2004.

[Dietrich 2006a]   D. Dietrich, P. Palensky, S. von Klot, D. Dietrich: „Digitale Gebäude", *Gebäudeinstallation - Fachmagazin für Sanitär, Lüftung, Klima, Enegrie und Wellness*, p.8-11, 2006.

[Dietrich 2007]   D. Dietrich, G. Fodor, W. Kastner, M. Ulieru: "Considering a technical realization of a neuro-psychoanalytic model of the mind", *Proceedings of 1st International Engineering and Neuro-Psychoanalysis Forum (ENF)*, p.13-20, 2007.

[DiPaola 2007]   S. DiPaola, C. Akai, B. Kraus: „Experiencing Belugas: Action Selection for Interactive Aquarium Exhibit"; *Adaptive Behavior*, Vol 15, No. 1, Sarge Publications, p. 99-113, 2007.

[Dodd 2005a]   R. Dodd, R. Gutierezz: "The Role of Episodic Memory and Emotion in a Cognitive Robot", Proceedings of 14th Annual IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN), Nashville, TN, p.692-697, 2005.

[Dodd 2005b]   W. Dodd: *The Design of Procedural, Semantic, and Episodic Memory Systems for a Cognitive Robot*, Thesis, submitted on the Faculty of Vanderbilt University, Nashville, Tennesee, 2005.

[Dreyfus 1989]   H.L. Dreyfus, S.E. Dreyfus: "Making a Mind Versus Modeling the Brain: Artifical Intelligence Back at a Branchpoint", *The Artificial Intelligence -*

	False Starts, Real Foundations, edited by Stephen R. Graubard, The MIT Press, p. 15-43, 1989.
[Egly 2004]	U. Egly, G. Novak, D. Weber, "Decision Making for Mirosot Soccer Playing Robots", *1st CLAWAR/EURON Workshop on Robots in Entertainment, Leisure and Hobby*, Vienna, Austria, Dec. 2-4, 2004.
[Fellous 1999]	J. M. Fellous, "Neuromodulatory basis of emotion", *The Neuroscientist*, Vol. 5, pp. 283-294, 1999.
[Fernandez 2001]	F. Fernández, L.E. Parker: "Learning in large cooperative multi-robot domains", *International Journal of Robotics and Automation*, 2001.
[Firby 1989]	R.J. Firby: *Adaptive Execution in Complex Dynamic Worlds*, Ph.D. thesis, Yale University, Computer Science Department. Technical Report 672.
[Fishbein 1975]	M. Fishbein, I. Ajzen: *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading*, MA: Addison-Wesley, 1975.
[Freud 1915]	S. Freud: „Triebe und Triebschicksale" (Instincts and their vicissitudes), *Internationale Zeitschrift für (ärztliche) Psychoanalyse*, III, p. 84-100; G.W., X, p. 210-232; S.E. 14: 117-140. available: „Triebe und Triebschicksale (1915)", *das Ich und das Es. Metapsychologische Schriften (Psychologie)*, Fischer Taschenbuch Verlag, 1911-1938, 11. unveränderte Auflage 2005, p.79-102, 1915.
[Freud 1917]	S. Freud: *Vorlesungen zur Einführung in die Psychoanalyse (1916-1917)*, Fischer Taschenbuchverlag, Copyright 1991, 13. unveränderte Auflage 2004.
[Freud 1923]	S. Freud: „Das Ich und das Es (1923), d*as Ich und das Es. Metapsychologische Schriften (Psychologie)*, Fischer Taschenbuch Verlag, 1911-1938, 11. unveränderte Auflage 2005, p.251-284, 1923.
[Fujimoto 2002]	Y.Fujimoto, T.Yakoh and K.Ohnishi: "Dynamic Model of Decentralized Systems With Informational Connection", *IEEE Transactions on Industrial Electronics*, Vol.49, No.3, p.707-715, 2002.
[Hearst 2000]	M. Hearst, H. Hirsh, "AI's greatest trends and controversies", *IEEE Intelligent Systems*, p.8-17, 2000.
[Henson 1998]	R.N.A. Henson: "Short-term Memory for Serial Order: The Star-End Model", *Cognitive Psychology 36*, Article No CG980685, p. 73-137, 1998.
[Holland 2001]	O. Holland: "From the Imitation of Life to Machine Consciousness", *Proceedings of International Symposium on Evolutionary Robotics From Intelligent Robotics to Artificial Life (ER 2001)*, p.1-37, 2001.
[IFR 2006]	IFR (International Federation of Robotics) Statistical Department: *World Robotics 2006: Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment*, 2006. online Version (02/2007): http://www.worldrobotics-online.org/

[ISO/TR 8373 1994] International Standard Organization: *Manipulating industrial robots – Vocabulary, reference number* ISO/TR 8373, first edition published 1994-12-0, 19941.

[Jones 2005] C.Jones, M. Maratić "Behavior-based Coordination in Multi-Robot Systems", *Autonomous Mobile Robots: sensing, Control, Decision-Making in Applications*, Sam S. Ge and Frank L .Lewis, eds. Marcel Dekker, Inc., 2005.

[Kabitzsch 2002] K. Kabitzsch, D. Dietrich, G. Pratl (Ed.), *LonWorks Gewerkeübergreifende Systeme*, VDE Verlag, 2002.

[Kant 1781] I. Kant: „Kritik der reinen Vernunft", Voltmedia, Paderborn, 1st edition in 1781, current edition in 2005.

[Kassabalidis 2001] I.Kassabalidis, M.A. El-Sharkawi, M.J. Marks, P. Arabshahi, A.A. Gray: "Swarm Intelligence for Routing in Communication Networks", *Proceedings of IEEE Global Telecommunications Conference*, p. 3613-3617 vol.6 2001.

[Kaplan-Solms 2000] K. Kaplan-Solms, M. Solms: *Clinical Studies in Neuro-Psycho-analysis*, International Universities Press, Inc., Madison, CT, 2000.

[Kastner 2004] W. Kastner, P. Palensky, T. Rausch, C. Roesener: "A Closer Look on Today's Home and Building Networks", *IEEE Africon*, p. 1239 – 1244, 2004.

[Kastner 2005] W. Kastner, G. Neugschwandtner, S. Soucek, M. H. Newman, "Communication Systems for Building Automation and Control", *Proceedings of the IEEE*, p. 1178-1203, 2005.

[Kitamura 1999] T. Kitamura: "Can a Robot's Adaptive Behavior be Animal-like without a Learning Algorithm?", Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, 1999 (IEEE SMC '99), Vol. II, p.1047-1051, 1999.

[Kitamura 2002] T. Kitamura, J.Otsubo, M. Abe: "Emotional Intelligence for Linking Symbolic Behaviors", Proceedings of IEEE Conference on Robotics & Automation, p. 1001-1006, 2002.

[Konolige 1996] K. Konolige, K. L. Myers, The Saphira Architecture for Autonomous Mobile Robots. in AI-based Mobile Robots: Case studies of successful robot systems, MIT Press, 1996.

[Kortenkamp 1998] D. Kortenkamp, R.P. Bonasso, R. Murphy: *Artificial intelligence and mobile robots: Case studies of successful robot systems*, MIT Press, 1998.

[Lerner 2000] J.S. Lerner, D. Keltner: "Beyond valence: Toward a model of emotion-specific influences on judgement and choice", *Cognition and Emotion*, Academic Press, Vol. 14, p.473-493, 2000.

[Lorenz 2007a]   B. Lorenz, E. Barnard: "A brief overview of artificial intelligence focusing on computational models of emotions", *Proceedings of 1ˢᵗ International Engineering and Neuro-Psychoanalysis Forum (ENF)*, p.1-12, 2007.

[Lorenz 2007b]   B. Lorenz: *A Neuro-Psychoanalytically Inspired Cognitive Architecture for Autonomous Systems,* PhD Thesis, Vienna University of Technology, to be published in 2007.

[Loy 2001]   D. Loy, D. Dietrich, H.-J. Schweinzer: *Open Control Network: Lon Works / EIA 709 Technology*, Kiuwer Academic Publisher, 2001.

[Lueg 1997]   C. Lueg, R. Pfeifer: "Cognition, Situatedness and Situated Design", *Proc. 2nd International Conference on Cognitive Technology*, p.124-135, 1997.

[MacKenzie 1997]   D. MacKenzie, R. Arkin, J. Cameron: "Multiagent mission specification and execution," *Autonomous Robots*, vol. 4(1), p. 29–52, 1997.

[Maes 1990]   P. Maes, R. Brooks: "Learning to Coordinate Behaviors", *Proceedings of Eighth national Conference on Artificial Intelligence (AAAI '90)*, p.115-120, 1990.

[Maes 1994]   P. Maes: "Modeling Adaptive Autonomous Agents", *Artificial Life: An Overview*, MIT press, p.135-162, 1994.

[Mahlknecht 2004a]   S. Mahlknecht, R. Oberhammer, G. Novak, "A Real-time Image Recognition System for Tiny Autonomous MobileRobots"; *Proceedings of the IEEE Real-Time and Embedded Technology and Applications*, pp. 324-330, 2004.

[Mahlknecht 2004b]   S. Mahlknecht: *Energy-Self-Sufficient Wireless Sensor Networks for Home and Building Environment*, PhD Thesis, Vienna University of Technology, 2004.

[McCarthy 2000]   J. McCarthy: *Concepts of logical AI*, paper in progress, Stanford University, available (11/2006): http://www-formal.stanford.edu/jmc/concepts-ai.html.

[McCarthy 2000a]   J. McCarthy: "Free will - even for robots" 3, J*ournal of Experimental and Theoretical Artificial Intelligence*, Vol. 12, Taylor & Francis, 2000.

[McCarthy 2002]   J. McCarthy: "Making Robots Conscious of their Mental States"16. In Stephen Muggleton, editor, *Machine Intelligence 15*, Oxford University Press, 1996-2002.

[McCarthy 2005]   J. McCarthy: "Simple Deterministic Free Will", invited talk on *The 7th International Symposium on Logical Formalizations of Commonsense Reasoning*, 2005.

available (10/2006): http://www-formal.stanford.edu/jmc/freewill2.html.

[McMullen 2002]   B. McMullen: "Cognitive Intelligence", *StudentBMJ (British Medical Journal)*, BMJ Publishing Group Limited, Volume 10, pp. 258-259, 2002.

[Meyer 1999]	W.-U. Meyer, A. Schützwohl, R. Reisenzein: *Einführung in die Emotionspsychologie, Band II, Evolutionspschologische Emotionstheorien*, 2. Auflage, Verlag Hans Huber, 1999.
[Milner 1989]	R. Milner: *Communication and Concurrency*, Prentice Hall, 1989.
[Milner 1999]	R.Milner: *Communicating and mobile systems: The π-calculus*, Cambridge University Press, 1999.
[Minsky 1986]	M. Minsky, *The Society in Mind*, Simon and Schuster, New York, 1986.
[Mochida 1995]	T. Mochida; A. Ishiguro, T. Aoki; Y. Uchikawa: "Behavior arbitration for autonomous mobile robots using emotion mechanisms", *Proceedings IEEE/RSJ Conf. on Intelligent Robots and Systems*, vol. 3, pp. 516-521, 1995.
[Mondada 1994]	R. Mondada, E. Franzi, and P. Ienne: "Mobile robot miniaturisation: A tool for investigation in control algorithms", *Proceedings of the Third International Symposium on Experimental Robots,* Springer, p.501-513, 1994.
[Mondada 1995]	F. Mondada, D. Floreano: "Evolution of Neural Control Structures: Some Experiments on Mobile Robots", *Robotics and Autonomous Systems, 16,* pp. 183-195, 1995.
[Novak 2002]	G. Novak: *Multi Agent Systems-Robot Soccer*, PhD Thesis, Vienna University of Technology, 2002.
[Novak 2004a]	G. Novak: "Roby-go, a Prototype for Serval MiroSOT Soccer Playing Robots, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 9, no. 4, p. 409 – 414, 2004.
[Novak 2004b]	G. Novak, A. Bais, S. Mahlknecht, "Simple Stereo Vision System for Real-Time Object Recognition for an Autonomous Mobile Robot", *Proceedings of the 2nd IEEE International Conference on Computational Cybernetics*, pp. 213-216, 2004.
[Novak 2005]	G. Novak and S. Mahlknecht: Tinyphoon - a tiny autonomous mobile robot, *IEEE ISIE 2005 International Symposium on IndustrialElectronics 2005*, p. 1533–1538, 2005.
[Novak 2006]	G. Novak, C. Roesener, M. Bader, T. Deutsch, S. Jakubek, S. Krywult, M. Seyr: "The Tinyphoon's Control Concept", *IEEE International Conference on Mechatronics*, 6p., 2006.
[Nuxoll 2004]	A. Nuxoll, J.E. Laird: "A Cognitive Model of Episodic Memory - Integrated With a General Cognitive Architecture", *Proceedings of the Sixth International Conference on Cognitive Modeling*, p. 220-225, 2004.
[Orebäck 2003]	A. Orebäck, H.I. Christensen: "Evaluation of architectures for mobile robotics", *Autonomous Robots 14*, 1 Kluwer Academic Publishers, p. 33-49, 2003.

[Orebäck 2004]   A. Orebäck, A Component framework for Autonomous Mobile Robots, doctoral thesis, KTH Numerical Analysis and Computer Science, Stockholm, 2004.

[Palensky 2003]   P. Palensky, P. Rössler, D. Dietrich: „Heim- und Gebäudeautomatisierung zur Effizienzsteigerung in Gebäuden", *Elektrotechnik und Informationstechnik (e&i)*, vol. 4, p. 104 – 110, 2003.

[Palensky 2005]   P. Palensky, D. Dietrich (Ed.), *Intelligente Gebäude und Wohnungen IGW+, planen, bauen, betreiben...,* Tagungsband der 5. IGW Konferenz, ÖVE, 2005.

[Palensky 2006a]   P. Palensky, „Intelligente Gebäude und Wohnungen - die Praxis", *Tagungsband der 6. IGW Konferenz*, Herausgabe einer Buchreihe, 2006.

[Palensky 2006b]   P. Palensky, S. Soucek, S. Von Klot, D. Dietrich, "Netzwerke und Gebäude," *Elektrotechnik und Informationstechnik* (e&i), vol. 6, pp. 259 – 268, 2006.

[Panait 2005]   L. Panait, S. Luke "Cooperative Multi-Agent Learning: The State of the Art", *Autonomous Agents and Multi-Agent Systems*, Volume 11, Number 3 / November, 2005.

[Panksepp 1998]   J. Panksepp: *Affective neuroscience: the foundations of human and animal emotions*, Oxford University Press, 1998.

[Panksepp 2002]   J. Panksepp, B. Knutson, J. Burgdorf: "The role of brain emotional systems in addictions: a neuro-evolutionary perspective and new 'self-report' animal model", Society for Study of Addiction to Alcohol and Other Drugs, p.459-469, 2002.

[Panksepp 2003]   J. Panksepp: "At the interface of affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain", *Brain and Cognition*, Academic Press, Vol. 52, p. 4-14, 2003.

[Panksepp 2004]   J. Panksepp: "Affective consciousness: core emotional feelings in animals and humans", article in press, 2004.

[Papert 1989]   S. Papert: "One AI or Many?", *The Artificial Intelligence - False Starts, Real Foundations*, edited by Stephen R. Graubard, The MIT Press, p. 1-14, 1989.

[Parker 1998]   L. E. Parker: ALLIANCE: An Architecture for Fault Tolerant Multirobot Cooperation, *IEEE Transactions on Robotics*, Vol. 14, No. 2, 1998.

[Parker 2002]   L. E. Parker: "Distributed Algorithms for Multi-Robot Observation of Multiple Moving Targets", *Autonomous Robots*, Volume 12, Number 3, p. 231-255(25), 2002.

[Parker 2006]   L. E. Parker, F. Tang: Building Multirobot Coalitions Through Automated Task Solution Synthesis, *Proceedings of IEEE*, Vol. 94, p.1289-1305, 2006.

[Pfeifer 1993]   R. Pfeifer: "Emotions in robot design", *Proceedings. 2nd IEEE International Workshop on Robot and Human Communication*, pp. 408 – 413, 1993.

[Pfeifer 1994]	R. Pfeifer: "The Fungus Eater approach to emotion: A view from artificial intelligence", *Cognitive Studies*, p.42-57, 1994.
[Pfeifer 1996]	R. Pfeifer: "Building Fungus Eaters: Design Principles of Autonomous Agents", *Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, vol. 4, From Animals to Animats, Cambridge, MA, p.10, 1996.
[Pfeifer 1999]	R. Pfeifer, C. Schreier, *Understanding Intelligence*, MIT Press, Cambridge MA, 1999.
[Pirjanian 1998]	P. Pirjanian: "Multiple Objective Action Selection in Behavior-Based Control", *The proceedings of the 6th Symposium for Intelligent Robotic Systems*, pp. 83-92, 1998.
[Plutchik 2001]	R. Plutchik: "The nature of emotions", *American Scientist*, vol. 89, no. 4: p.344-350. 15, 2001.
[Prassler 2000]	E. Prassler, A. Ritter, C. Schaefer, P. Fiorini: "A Short History of Cleaning Robots", *Autonomous Robots 9*, Kluwer Academic Publishers, p. 211-226, 2000.
[Pratl 2005a]	G. Pratl, W.T. Penzhorn, D. Dietrich, W. Burgstaller: „Perceptive Awareness in Building Automation", *Proceedings of the ICCC 2005 IEEE 3rd International Conference on Computational Cybernetics*, p. 259-264, 2005.
[Pratl 2005b]	G. Pratl, B. Lorenz, D. Dietrich: "The Artificial Recognition System (ARS): New Concepts for Building Automation", *Proceedings of 6th IFAC International Conference on Fieldbus Systems and their Applications*, p. 48 – 55, 2005.
[Pratl 2006]	G. Pratl, *Processing and Symbolization of Ambient Sensor Data*, PhD Thesis, Vienna University of Technology, 2006.
[Pratl 2007]	G. Pratl, D. Dietrich, G. P. Hancke, W. T. Penzhorn, W. "A New Model for Autonomous, Networked Control Systems", *IEEE Transactions of Industrial Informatics*, Digital Object Identifier 10.1109/TII.2007.891308, pp.21-32, 2007.
[Prescott 2007]	T.J. Prescott: "Forced Moves or Good Tricks in Design Space? – Landmarks in the Evolution of Neural Mechanisms for Action Selection", *Adaptive Behavior*, Vol 15, No. 1, Sarge Publications, p. 9-31, 2007.
[Ramachandran 2004]	V.S. Ramachandran, S. Blakeslee: *Phantoms in the brain: probing the mysteries of the human mind*, Quill ed., 2004.
[Ratanaswasd 2005]	P. Ratanaswasd, S. Gordon, W. Dodd: Cognitive control for robot task execution, *IEEE International Workshop on Robot and Human Interactive Communication*, ROMAN 2005, pp. 440 – 445, 2005.

| [Roesener 2004] | C. Roesener, H. Hareter, W. Burgstaller, G. Pratl: "Environment Simulation for Scenario Perception Models", *5th IEEE International Workshop on Factory Communication System*, p. 349 – 352, 2004. |

[Roesener 2006]   C. Roesener, B. Lorenz, G. Fodor, K. Vock: "Emotional Behavior Arbitration for Automation and Robotic Systems", *4th International IEEE Conference on Industrial Informatics*, INDIN´06, 6p, 2006.

[Roesener 2007]   C. Roesener, R. Lang, T. Deutsch, A. Gruber, B. Palensky: "Action planning model for autonomous mobile robots", *5th International IEEE Conference on Industrial Informatics*, INDIN´07, p.6, 2007, to be published.

[Rupp 2005]   C. Rupp, J. Hahn, S. Queins, M. Jeckle, B. Zengler: "UML 2 glasklar", Carl Hanser Verlag, 2005.

[Russ 2003]   G. Russ, *Situation Dependent Behavior in Building Automation*, PhD Thesis, Vienna University of Technology, 2003.

[Sacks 1986]   O. Sacks: *The man who mistook his wife for a hat*, Pan Books, 1986.

[Schraft 2004]   R.D. Schraft, M. Hägele, K. Wegner: Service Roboter Visionen, Carl Hanser Verlag, 2004.

[Shadbolt 2003]   N. Shadtbolt: „In Memoriam", *IEEE Intelligent Systems*, p.2-3, 2003.

[Shell 2005]   D.A. Shell, M.J. Matarić: "Behavior-Based Methods for Modeling and Structuring Control of Social Robots", *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Ron Sun, ed., Cambridge University Press, pp. 279-306, 2005.

[Shibata 1996]   T. Shibata, K. Ohkawa, K. Tanie: "Spontaneous behavior of robots for cooperation: Emotionally intelligent robot system", *IEEE Proceedings on Int. Conf. on Robotics and Automation*, vol. 3, pp. 2426-2431, 1996.

[Shoham 1993]   Y. Shoham: "Agent Oriented Programming", *Journal of Artificial Intelligence* 60 (1), pp. 51-92, 1993.

[Siegwart 2004]   R.Siegwart, I.R.Nourbakhsh: *Introduction to Autonomous Mobile Robots*, The MIT Press, 2004.

[Singh 2003]   P. Singh; M. Minsky: "An architecture for combining ways to think". *Proceedings Int. Conf. on Knowledge Intensive Multi-Agent Systems,* Cambridge, MA.; pp. 669-674, 2003.

[Sloman 2002]   A. Sloman. "How many separately evolved emotional beasties live within us?". *Emotions in Humans and Artifacts* (R. Trappl et al., eds), MIT Press, p.35-114, 2002.

[Solms 1997]   M. Solms: "What is Consciousness?", *Journal of the American Psychoanalytic Association*, vol. 45, pp. 681-778, 1997.

[Solms 2002]   M. Solms, O. Turnbull: *The Brain and the Inner World*, Karnac/Other Press, Cathy Miller Foreign Rights Agency, London, England, 2002.

[Stewart 2005]   T. Stewart: *Two cognitive Descriptions of Q-Learning,* Carleton University Cognitive Science Technical Report 2005-03, 2005.

[Stichweh 2004]   R. Stichweh: *Automaten. Über Norbert Wiener, "Cybernetics or Control and Communication in the Animal and the Machine" (1948)*, essay 2004.

available (01/2007): http://www.unilu.ch/files/18stwwiener.pdf

[Sugawara 1999]   K. Sugawara, M. Sano, I. Yoshihara, K. Abe, T. Watanabe: "Foraging Behavior of Multi-robot System and Emergence of Swarm Intelligence", *Proceedings. 1999 IEEE International Conference on IEEE SMC '99 Conference*, Volume: 3, p. 257-262, 1999.

[Taipale 1993]   T. Taipale, S. Hirai: "A behavior-based control system applied over multi-robot system", *Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems '93*, IROS '93, pp. 1941 - 1943 vol.3, 1993.

[Tamarit 2003]   C. Tamarit, *Automafion System Percepfion,* PhD Thesis, Vienna University of Technology, 2003.

[Taylor 2006]   R.T. Taylor: "A Perspective on Medical Robots", Proceedings of the IEEE, Vol.94, No.9, p.1652-1664, 2006.

[Toda 1982]   M. Toda: *Man, Robot and Society*; Martinus Nijhoff Publishing, 1982.

[Tulving 1983]   E. Tulving: *Elements of episodic memory*, Cambridge University Press, 1983.

[Turkle 1989]   S. Turkle: "Artificial Intelligence and Psychoanalysis", *The Artificial Intelligence - False Starts, Real Foundations*, edited by Stephen R. Graubard, The MIT Press, p. 241-268, 1989.

[Turing 1950]   A. M. Turing: „Computing Machinery and Intelligence", *Mind*, 59, p.433-460, 1950.

[Uchimura 2002]   Y. Uchimura, K. Ohnishi and T. Yakoh: "Bilateral robot system on the real time network structure", *IEEE International Workshop on Advanced Motion Control (AMC-2002)*, p.63-68, 2002.

[Uchimura 2004]   Y.Uchimura, K.Ohnishi and T.Yakoh, "A controller design method on decentralized systems", *Industrial Electronics Society, 2004. 30th Annual Conference of IEEE (IECON-2004)*, Vol.3 p.2961-2966, 2004.

[Ueno 1999]   A. Ueno, H. Takeda, T. Nishida: "Cooperation of Cognitive Learning and Behavior Learning"*, Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, p.387-392, 1999.

[Ueno 2000]   K. Uneo, T. Yakoh, T. Murakami, K. Ohnishi: "New Representation of Systems with effect Vairables", *Proceedings of IEEE 6th International Workshop on Advanced Motion Control*, p. 216-221, 2000.

[Umeda 2002]   Y. Umeda, T. Yakoh: "Configuration and Readhesion Control for a Mobile Robot with External Sensors", *IEEE Transactions on Industrial Electronics*, Vol.49, No.1, p.241-247, 2002.

[UNECE 2004]   United Nations Economic Comission for Europe (UNECE): *World Robotics 2004: Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment*, United Nations Publications, press release, 2004.

[UNECE 2005]   United Nations Economic Comission for Europe (UNECE): *World Robotics 2005: Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment*, United Nations Publications, executive summary, 2005.

[Wagner 2002]   A.D. Wagner. "Cognitive Control and Episodic Memory", *Neuropsychology of Memory*, by editors L.R. Squire, D. L. Schacter, Third Edition, The Guilford Press, p. 174-192, 2002.

[Wehrle 1994]   T. Wehrle: "New fungus eater experiments (autonomous agents)", *Proceedings of from Perception to Action Conference,* P. Gaussier, & J.-D. Nicoud (Eds.), Los Alamitos: IEEE Computer Society Press, pp. 400-403, 1994.

[Weiser 1991]   M. Weiser: "The Computer for the 21st Century", *Scientific American 265 (1991)*, Nr. 3, p. 66–75, 1991.

[Whiteson 2007]   S. Whiteson, M.E.Taylor, P. Stone: "Empirical Studies in Action Selection with Reinforcement Learning", *Adaptive Behavior*, Vol 15, No. 1, Sarge Publications, p. 33-50, 2007.

[Yakoh 1993]   T. Yakoh, Y.Anzai: "Physical Ownership and Task Reallocation for Multiple Robots with Heterogeneous Goals", *Proceedings of International Conference on Intelligent and Cooperative Information Systems*, IEEE, p.80-87, 1993.

Internet references

[eAnylogic 2007]   Homepage of the project AnyLogic

Available (06/2007): http://marie.sourceforge.net/

[eARS 2007]   The Artificial Recognition System, homepage of the project, Institute of Computer Technology, Vienna University of Technology

Available (05/2007): http://ars.ict.tuwien.ac.at/

[eBACM 2007]   Technical Committee of the Building Automation, Control and Management (TC BACM), Industrial Electronics Society (IES), IEEE

Available (04/24/2007):http://www.ies-bacm.org/

[eBahill 2007]   A. T. Bahill: Online library of publications and definition

Available (01/24/2006): http://gd.tuwien.ac.at/systeng/bahill/definitions.htm

[eCapek 2007]   Biography of Karel Capek

Available (01/2007): http://home.clara.net/heureka/art/karel-capek.htm

[eMarie 2007]   Homepage of the project MARIE (Mobile and Autonomous Robotic Integration Environment)

Available (06/2007): http://marie.sourceforge.net/

[ePyro 2007]   Homepage of Pyro, a Python-based Programming Environment

Availabale (06/2007): http://www.pyrorobotics.org/

[eStanford 2007]   Stanford Encyclopedia of Philosophy: Emotion

Available (05/2007): http://plato.stanford.edu/entries/emotion/

CURRICULUM VITAE

Charlotte Rösener has studied at Vienna University of Technology Electrical Engineering and Computer Science. In November 2002 she graduated with distinction. The objective of her diploma thesis are special aspects of fault tolerance, in particular defect tracking systems in circuitry, network management, and fault tolerance in the human body (bionics). Between May 2003 and March 2005 she worked as a scientific employee at the Institute of Computer Technology (ICT) of Professor PhD Dietmar Dietrich. The main focus on her scientific work was intelligent networks in building automation, autonomous behavior arbitration of mobile robots, and information technology for energy networks. During her employment on the Institute of Computer Technology (ICT) she participated in the projects JEVis, IRON et al. In April 2005 she entered a program to work as a research student in Professor PhD. Takahiro Yakoh's Yakoh Laboratory of, Department of System Design Engineering, Faculty of Science and Technology, Keio University, Tôkyô, Japan. Furthermore she also participates in KNIVES project of Nishi Laboratory of Professor PhD. Hiroaki Nishi. During this period she worked simultaneously on her PhD on Institute of Computer Technology (ICT).

Education

09/87 – 06/91	AHS Wenzgasse 7, A-1130 Vienna (high school)
09/91 – 06/95	ORG Hegelgasse 14, A-1010 Vienna (high school)
12.06.1995	school leaving examination passed with distinction

Studies

10/95 – 11/02	Electrical Engineering on Vienna University of Technology, Austria
	Major: Computer Technology,
	Diploma examination passed with distinction
	Diploma thesis *„Fallbeispiele und Aspekte fehlertoleranter Systeme aufbereitet für das Internet"* – Institute of Computer Technology
	(adviser: O. Univ. Prof., Dr. tech., Dipl.Ing. Dietmar Dietrich)
2000	CIM-Project (BMW – Univ. Prof. Dr. techn. Gerfried Zeichen)
2001	Web-based Publication Database (O.Univ. Prof. Dr. techn. Karl Riedling)
10/03 – 05/07	„Doctorate-studies of science", Field of dissertation: Electrical Engineering
10/03 – 03/05	Japanese studies on University Vienna, Austria
04/05 – 03/07	Research student (MEXT scholarship) on Keio University, Japan

Activities during studies

07/97 – 09/97	practical training in Siemens Inc., 1100 Vienna, software engineering
07/99 – 09/99	practical training in Spardat Ltd., 1110 Vienna, software engineering

Professional experience

10/99 – 05/03	Software engineer, department "Intra/ Internet", Spardat (Sparkassen Datendienst) Ltd., Vienna
05/03 – 03/05	Employed as project assistant, Institute of Computer Technology, ICT (384), Vienna University of Technology

Committees

09/04 – 03/05	Thematic Network "Energie und Kommunikation" im Vernetzungsfond erneuerbare Energien, Germany
01/06 – to date	IEEE students member
01/06 – to date	IEEE Industrial Electronics Society (IES)
01/06 – to date	Technical Committee of the Building Automation, Control and Management (TC BACM)