Brain-Like Artificial Intelligence for Automation – Foundations, Concepts and Implementation Examples

Rosemarie Velik
CTR Carinthian Tech Research, Villach, Austria
E-mail: Rosemarie.velik@ctr.at

Abstract:

Over the last decades, automation technology has made serious progress and can today automate a wide range of tasks having before needed human physical and mental abilities. Nevertheless, a number of important problem domains remain that cannot yet be handled by our current machines and computers. A few prominent examples are applications involving "real-world" perception, situation assessment, and decision-making tasks. Recently, researchers have suggested to use concepts of "Brain-Like Artificial Intelligence", i.e. concepts inspired by the functioning principles of the human or animal brain, to further advance in these problem domains. This article discusses the potential of Brain-Like Artificial Intelligence for innovative automation solutions and reviews a number of approaches developed together with the ICT cognitive automation group of the Vienna University of Technology targeting the topics "real-world" perception, situation assessment, and decision-making for applications in building automation environments and autonomous agents. Additionally, it is demonstrated by a concrete example how such developments can also contribute to an advancement of the state of the art in the field of brain sciences.

Keywords: Brain-like artificial intelligence, cognitive automation, machine perception, recognition, situation assessment, decision-making,

1. Introduction

Automation has made considerable progress in observing and controlling processes of various kinds over the last decades. A discipline that has substantially contributed to these advancements is the domain of Artificial Intelligence (AI). Nevertheless, the successes we achieved by automation and AI are often still limited to relatively specific, well-definable tasks in well-structured environments (e.g., the usual case in production lines of factories). When we shift to applications where a broader range of tasks should be performed in less organized environments (e.g., the automation of processes in offices or private homes that need to incorporate various activities of their occupants into their control strategies), we still face challenges that are often beyond today's technical feasibility. In such "multi-faceted real world ambiences", our machines are to date frequently not able to correctly recognize the large number of possible occurring scenarios and situations, to assess and classify them (as harmless or harmful, beneficial or disadvantageous, etc.), and to decide about the adequate measures and (re-)actions corresponding to the given situation. It is doubtful if mere iterative and incremental improvements of today's approaches and concepts to automation and AI will ever be able to lead us to such a breakthrough; a paradigm shift might be necessary to succeed.

In contrast to current machines, humans (and also higher developed animals) can cope very well with such environments. Their brain reconstructs the environment from the incoming stream of (often ambiguous) sensor information, generates unambiguous interpretations of the world on a more abstract level, evaluates these perceptions and concepts, and makes adequate decisions about how to (re-)act to them. Thus, developing machines based on the same concepts and information processing principles as the human brain has considerable potential for applications in various domains of automation. The research field of "Brain-Inspired" or "Brain-Like Artificial Intelligence" [37, 49] has taken up the challenge of deciphering the working mechanisms of the brain and translating them into technically implementable concepts.

Brain-Like Artificial Intelligence is still a relatively recent and dynamic domain. In [61], we gave a general review about the state of Brain-Like Artificial Intelligence and outlined its potentials and current flaws. In contrast to this general overview, the article at hand aims at reviewing the potentials, challenges, and particularities of this young and dynamic research domain in the context of automation technology. While Chapter 2 and 3 are concerned with a discussion of this topic on a general level, Chapter 4 provides concrete implementation examples by reporting about own research work. In this context, novel brain-inspired architectures for complex robust machine perception, situation assessment, and decision-making are presented and the feasibility of their application for automation is demonstrated and discussed. Furthermore, it is shown by a concrete example how research insights gained during such brain-inspired development processes can contribute to valuable new hypotheses concerning brain functioning.

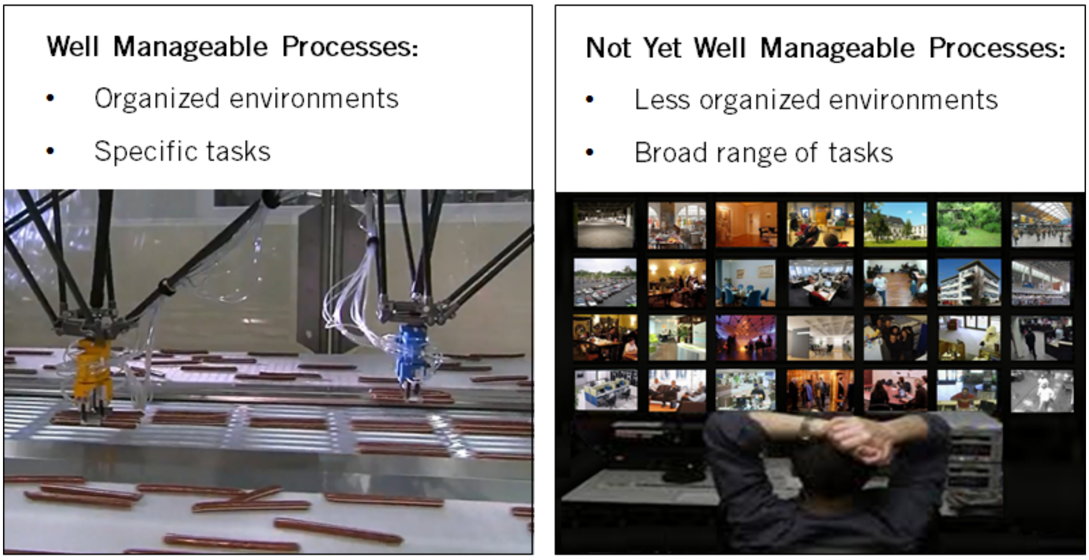
2. The Need for a Paradigm Shift in Certain Areas of Automation

2.1 General Objectives of Automation

The term automation stems from the Greek word "automatos" meaning self-moving, self-thinking, or self-acting. The field of automation is concerned with the employment of scientific and technological principles for the design and construction of machines that take over tasks previously performed by humans. Automation has its origins in the mechanization of processes, i.e., the provision of machinery tools to assist human operators in the physical requirements of work. However, within the last decades, significant progress has been made in this field and automation has developed beyond pure mechanization and is today able to additionally decrease the need for human sensory and cognitive abilities or to complement them by electronic and computational means. The general objective of today's automation is to reduce the need for human (physical as well as cognitive) work force in the production of goods and the provision of services. This has traditionally been postulated to be particularly advantageous in (1) tasks that involve hard physical work, (2) tasks that are monotonous, (3) tasks performed in dangerous environments, and (4) tasks that are on the edge of or beyond human capabilities of size, weight, speed, accuracy, or endurance. Apart from applications in production and construction environments, this can also include operations such as rescue work, space missions, or exploration of ocean floors. Furthermore, automation has started to penetrate our daily living environments. Solutions already put into place are, e.g., building automation systems for energy management, access control, fire and hazard detection. Potentials for further applications in such "everyday life domains" are huge, including amongst others services for increased comfort and entertainment and services for people who need them would prefer to have done by some kind of machine instead of other human beings (e.g., care for disabled persons, household tasks, etc.). A domain that has significantly contributed to the progress in automation is the field of Artificial Intelligence (AI) [2]. Nowadays, automation technology incorporating AI methods is spreading over a wide range of applications beyond manufacturing including applications like banking systems, plant control, traffic control, flight control, power distribution management, information transmission, code generation, and many more. As AI advances, improvements in automation and novel target applications will follow.

2.2 Current Limitations

Despite the many progresses that have been achieved in automation and AI over the last decades, applications have mainly been successful for circumscribed, well-defined tasks in environments that are relatively specific, well-structured, and characterized by a limited number of possible occurring processes, states, and ways how to react to them [59]. There, technical solutions can even exceed certain human capabilities. The situation changes however if we switch to systems that should perform a broader range of tasks in less well-structured environments. Here, the limits of technical feasibility are being stretched to the utmost [7]. This fact is probably best illustrated by the following two concrete examples in Figure 1.



(a) Industrial Sorting Task (b) Safety and Security Surveillance Task

Figure 1. Examples for Today Well Manageable and not yet Well Manageable Processes in Automation

The first example given in Figure 1a shows an industrial sorting task, specifically the sorting of salami sausages into two classes according to their size. With current automation technology, such processes can be handled relatively well because the framework conditions are relatively simple and clear. All occurring objects are per definition salami sausages on a conveyor belt and the task to be performed is their separation into two categories based on their size.

An example of a task of a completely different level of complexity can be found in Figure 1b, which refers to the safety and security surveillance of a public building. As indicated in the picture, in order to achieve high system reliability and a low number of false alarms, which can be extremely costly, such applications usually still have to rely on the perceptual and cognitive abilities of human surveillance personnel. To date, machines are not able to correctly recognize the large number of possible occurring scenarios and situations, to correctly classify them as harmless, suspicious, safety-critical, or harmful, and to decide which are the adequate measures and corresponding (re-)actions to take according to the given situation.

Generalizing from the two examples presented above, the following observation can be made: no matter what application domain, developed systems and algorithms are to date usually designed for isolated, well-structured tasks of limited complexity. Almost all automation systems are reactive in their operation; i.e., these systems classify states and react to them based on relatively “simple”, explicit, pre-defined rules or mathematical constructs [42, 43]. It is hardly possible to integrate them into a larger, more capable all-purpose system [14]. However, with the shift in application to more “natural” environments, the requirements and demands put on automation systems have started to change in the last years. A number of envisioned environments – particularly daily living environments – have shown not to follow simple and “deterministic” laws as soon as tasks become more “demanding” than, e.g., today's “reactive” lighting and building climate control (heating, ventilation, air conditioning) [5, 31, 32, 35]. For more “sophisticated” applications in more “natural” environments including for instance the detection and interpretation of ongoing human activity, the number of possible occurring scenarios and the way how to optimally respond to them is seemingly infinite [24, 25, 50, 59]. This severely challenges the approaches available today. Apart from the field of safety and security surveillance described in the example from Figure 1b, similar challenges are present in other recently envisioned application domains such as the followings:

- In the domain of energy and resource management of public, commercial, and private buildings, the control of building climate and lighting currently depends on relatively “simple” information coming from occupancy sensors, thermostats, timers, and the like. What would however be of additional value for energy and resource saving and carbon footprint optimization is the consideration of the overall situation of ongoing and anticipated (occupant) activities in the building and the corresponding decisions of how to efficiently regulate resources accordingly without decreasing the comfort level of the inhabitants or the quality of provided services [59].

- One of the main goals of telecare is to allow the elderly and physiologically or mentally disabled persons to stay longer in their own homes and live more independently. This is achieved by monitoring persons' activities and detecting critical or harmful situations (e.g., falls) to inform family members or care attendants in case of their occurrence. However, similar as in the field of safety and security surveillance, current technologies show only limited trustworthiness and still give rise to a considerable number of false alarms. Accordingly, they are not yet widely employed [5, 6, 8, 9, 32]. More sophisticated concepts are therefore desirable.
- An improvement of concepts would also be advantageous in the area of interactive systems in order to offer new ways of user interaction and increase the speed of interaction, satisfaction, and comfort of users by interpreting their intentions and accordingly adapting to their needs [53].
- In the future, flexible and robust autonomous robots and other agents could be employed for assistance in homes, offices, public spaces, and factories, for work in hazardous or unpleasant environments, and for space and ocean exploration. However, existing mechanisms for robot perception, action selection, and action execution are in many situations still not sufficiently flexible and fault-tolerant to allow for a "smooth" autonomous and sensible navigation, object manipulation, and interaction in such complex environments [12, 46, 56]. Novel, more capable concepts are needed here as well.

2.3 Three Important Challenges to Face

The list of desirable but yet not satisfactorily automatable applications given in Section 2.2 can easily be extended. In principle, all processes and tasks can be put on this list that today still need to be performed by humans due to a lack of machine capabilities to handle them. Three issues that should particularly be addressed to allow for future progress in such application domains are the following:

- More advanced methods of **machine recognition** are needed, i.e., novel concepts of sensor data processing, sensor fusion, and data mining, most likely based on information from a larger number of sensors of different, partly redundant and partly complementary sources in order to unambiguously recognize relevant objects, events, and activities in complex environments [58].
- New concepts for **situation assessment** are necessary to evaluate the meaning that these objects, events, and activities have in a current or future context [28, 34].
- Adequate (re-)actions should occur based on the made recognitions and evaluations, e.g., in the form of some alarm, alert, or control of certain actuators. In a part of the envisioned future applications, the triggering of the "corresponding" (re-)actions is quite straightforward (e.g., occupant-activity-dependent lighting and building climate control); however, in other cases, the best action has to be selected from a large number of possibilities having different short and long-term benefits and drawbacks (see [59]). In such complex environments, the outcomes and consequences of taken actions cannot always be predicted with certainty. Such applications therefore require more capable **decision-making** mechanisms [16].

2.4 Two Possible Routes to Proceed

Having identified the need for novel methods for machine recognition, situation assessment, and decision making in order to advance further in different automation domains, an important question is by what means can we reach such sophisticated mechanisms. The long-term goal in mind is to construct machines and systems showing performances comparable to or even beyond human skill levels. In a guest talk at the Vienna University of Technology in 2008, Prof. Etienne Barnard, an expert in the field of Artificial Intelligence, made an interesting "conceptual suggestion" for two possible progress scenarios to reach this goal which could be summarized as depicted in Figure 2.

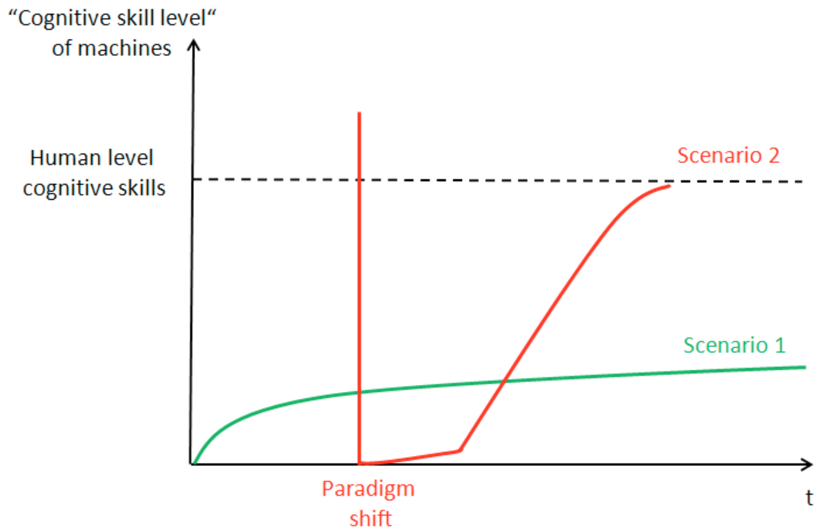


Figure 2. Two Possible Progress Scenarios for How to Reach Towards Machines and Systems with Human-Level Cognitive Skills

The first possibility for how progress could be achieved in the future is indicated in Scenario 1. According to this scenario, we continue on the current strategic and technological path and improve approaches and concepts in automation and artificial intelligence incrementally over time by small step-by-step innovations. In this way, we will eventually one day reach performances comparable to human skill levels.

An alternative to this incremental path is proposed in Scenario 2. Scenario 2 suggests that instead of waiting for incremental improvements over a very long time horizon, which could also stagnate at a state that is by far below human skill levels, a paradigm shift in strategic approaches, concepts, and/or technologies could occur. After an initial phase of research, this paradigm shift could then lead to synergetic effects making it possible to achieve the targeted goal of reaching human skill levels in machines much faster than in Scenario 1.

Looking at Figure 2, Scenario 2 definitely seems attractive. The question that remains to answer is of course through what revolutionary idea(s) could such a paradigm shift be introduced. In this article, one potential way for initiating such a paradigm shift in certain fields of automation is proposed. The solution of choice is to develop concepts of "Brain-Inspired" or "Brain-Like Artificial Intelligence" for complex recognition, situation assessment, and decision-making tasks. The principal feasibility of this approach is investigated and tested by introducing various Brain-Inspired AI architectures, implementing them, applying them to different automation tasks, and analyzing the results and insights gained (see Chapter 4).

3. Foundations of Brain-Like Artificial Intelligence for Automation

3.1 Basic Structure of Artificial and Biological (Brain-Controlled) Automation Systems

Although basing on different concepts concerning their details, artificial and biological (brain-controlled) automation systems show common points concerning their principal components (see Figure 3).

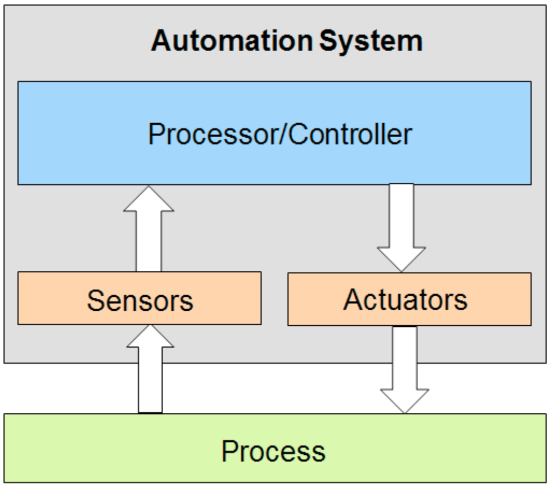


Figure 3. Basic "Components" of Artificial and Biological (Brain-Controlled) Automation Systems

In both cases, the starting point is a "process" that shall be controlled. From this process, selected process variables have to be observed via appropriate sensors – in one case technical sensors, in the other case biological sensors in form of physiological receptors. The sensory information is then further processed via a technical processor/controller and the brain/nervous system, respectively. Afterwards, the process variables are adapted by taking influence on the process via appropriate actuators (in the case of artificial systems, e.g., via the control of motors, in biological systems, e.g., via the the control of muscles). Resulting changes in the process variables can then again be observed via the sensors and a new sensing-processing-actuating cycle is started.

3.2 Basic Idea and Motivation of Brain-Like Artificial Intelligence

The field of Brain-Like Artificial Intelligence as defined in this article is relatively recent and dynamic. It can be considered as belonging to the broader group of bio-inspired technical approaches. As a matter of fact, biological systems and principles have already proven in the past to be valuable sources for technological development. One prominent discipline building on inspiration from biological principles for studying and designing engineering systems is the field of bionics [39]. Here, approaches consider, e.g., the construction of flying machines by studying the flight of birds or the functioning principle of biological receptors for the design of innovative sensors. A further related discipline is the field of cybernetics, which is, amongst others, concerned with the studying of control and communication mechanisms as well as concepts in animals and machines [65]. Another field that could be considered as the most direct ancestor of Brain-Like Artificial Intelligence is the domain of neural networks [48]. Here, the study of the functioning principles of individual neurons led to mathematical models applicable to certain pattern recognition, function approximation, and prognosis tasks.

The principal idea and motivation behind the field of Brain-Like Artificial Intelligence as followed in this article is quite straightforward: As outlined in the examples from Figure 1, the dealing with a broad range of tasks in complex environments today still needs human perceptual and cognitive skills. The only system currently successful in processing such multifaceted information is thus the (human) brain. It is unclear, even doubtful, whether a mere further development of the currently employed automation and AI paradigms will be able to change this fact in the coming decades or even century. Therefore, to approach the automation of such challenging tasks, a promising alternative to current (mainly purely mathematic/algorithmic) automation and AI concepts is to investigate in more detail how the brain manages to solve these tasks and to then take over these concepts for the development of technical systems. Evolution has equipped our brains with highly efficient circuits and mechanisms for processing sensory information gathered from millions of sensory receptors, evaluating this information, and making decisions despite numerous possibilities, contradicting aims, and uncertain outcome. Deciphering these circuits, mechanisms, and functions on the neural and/or cognitive level and translating them into technically implementable concepts, as attempted in Chapter 4 of this article, could have the potential to lead to a revolution in certain fields of automation and machine intelligence and to bring about major scientific advances together with social and economic benefits [15, 45, 58].

The research field of Brain-Like Artificial Intelligence for automation is still relatively novel and dynamic and breaks with some established dogmas in automation, AI, and engineering in general. Accordingly, differences exist concerning the way how research is/should be performed in comparison to more conventional approaches in engineering and computer sciences. Two significant differences lie in the applied scientific methodology and the validation criteria. As these two points often lead to misconceptions and debates when not outlined explicitly – and are thus particularly critical for the judgment of research in this field – the following two sections describe the methodology followed in this work and the validation criteria considered appropriate for judging it.

3.3 Methodology of Brain-Like Artificial Intelligence

An overview of the methodology for developing Brain-Like AI architectures for automation – as applied in the research of this article – is sketched in Figure 4 and further described in the following enumeration:

1. The starting point for development is, similar as in other domains of automation, a given automation problem and an identification of the requirements.
2. In the classical domain of automation, the next step would now be the elaboration of different potential approaches to solution for the given problem and their comparison. This second step already constitutes the first difference between the field of Brain-Like AI and classical automation. As in Brain-Like AI, brain-inspired concepts are employed, the next step after identifying the automation problem is to evaluate the brain sciences with the aim to determine – as far as known – how the brain manages to solve the given task.
3. Having identified the adequate processing concepts, the next step is the derivation of a technically implementable model based upon these insights. Performing step two and step three is of course far from trivial. A relatively broad body of knowledge is provided from brain sciences concerning the neural level and the functional level. However, between those two levels, a gap exists in understanding how neural activity correlates with cognitive function. The challenge to face in Brain-Like AI for automation is therefore to close this gap in an adequate way in the technical models. (See Chapter 4 for examples of models that are attempting to solve this challenge.)
4. After having developed and implemented the model, the next step is the validation of the model concerning its performed function. In a first instance, this is usually achieved via computer simulations. Particularities concerning this validation process are outlined further in Section 3.4.
5. The next step would then be the development of a demonstrator or – more advanced – the design of the automation system. In certain cases, step 4 an step 5 can also be merged.
6. As a side effect – besides their utility for automation systems – the developed models can furthermore lead to the formulation of new hypotheses concerning brain functioning and therefore contribute to the body of knowledge in brain sciences.

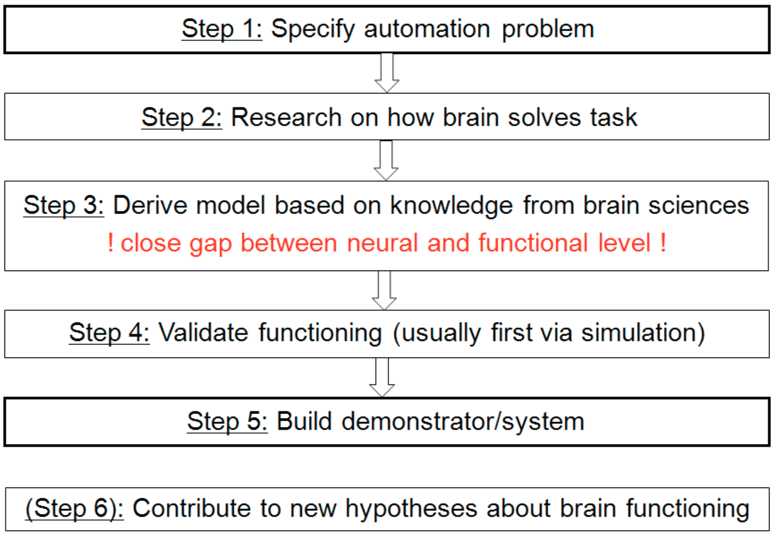


Figure 4. Methodology in the Field of Brain-Like Artificial Intelligence for Automation

3.4 Validation Criteria of Brain-Like Artificial Intelligence

A usual validation procedure in classical fields of engineering and computer sciences as well as in Applied AI, which is currently the dominant AI research domain, is to analyze and implement different potential methods to solve a given problem and to then compare their performance. What is thus usually desired are comparable, quantifiable results. In comparison, the starting situation is different in the field of Brain-Like AI (see Figure 5).

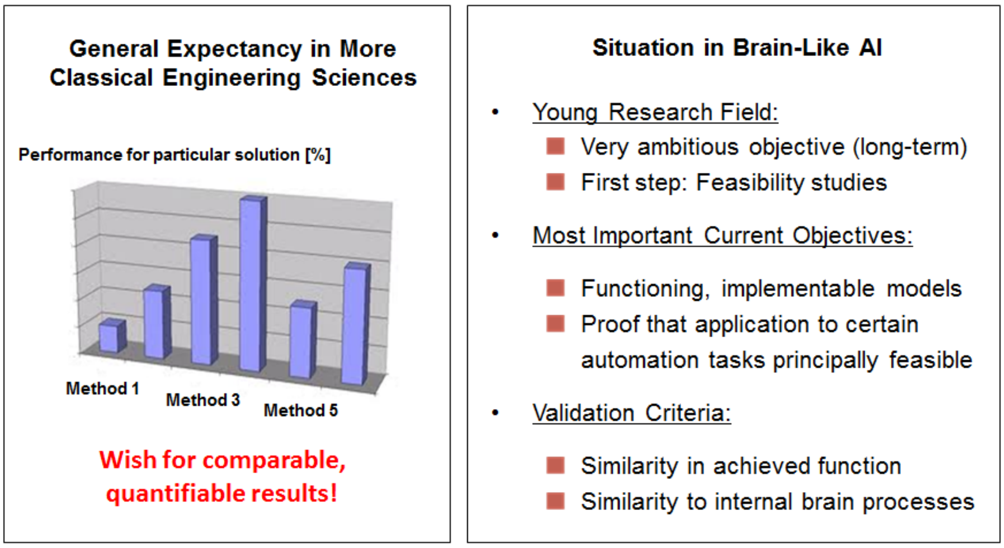


Figure 5. Differences in Validation Criteria between Classical Engineering Sciences and the Field of Brain-Like Artificial Intelligence for Automation

After all, Brain-Like AI is still a relatively young research field with very ambitious objectives that can only be targeted in the long run. Unlike in other fields of engineering, the goal is not only to solve a given task (by whatever suitable technical approach), but also to take a look how the brain solves this task and to take over these processing principles and concepts. Thus, Brain-Like AI is currently facing a much higher level of complexity in developing a solution. The vision is that in the long run this will lead to greater benefits and improvements than today's more “conservative“ approaches. Today, we are at the very beginning of this process and still face many stumbling blocks. Before being able to aim at competition and optimization, the first much more important hurdle to overcome is to validate the feasibility of developing functioning brain-inspired methods and applying them for a certain range of automation tasks. These selected applications and tasks constitute challenges that cannot yet be solved satisfactorily by existing technical approaches.

Accordingly, the aim of this article is to provide feasibility studies by developing different brain-inspired models, implementing them, and employing them to different challenges and applications. To conclude if an approach has been successful in achieving this objective, it can be validated as follows.

The goal of research in Brain-Like AI is to emulate information processing concepts of the brain for different tasks and applications. Therefore, to validate the value of a model in such a feasibility study, it should actually not be compared against other technical systems but against the brain. This can be done using the following two (currently rather qualitative than quantitative) criteria, which should ideally be fulfilled as much as possible:

1. **Similarity in Achieved Function:** How well does the developed and implemented model solve the chosen task in comparison to the brain?[1]
2. **Similarity to Internal Brain Processes:** To what extent do the structural organization and the information processing concepts of the developed model show similarities to those of the brain?[2]

The importance of fulfilling not just one but both of the abovementioned criteria can be illustrated by the following simple example:

Assume that the chosen task is to perform simple addition and subtraction problems of four-digit numbers. This function could easily be achieved by common logic circuits, which could solve this task even faster and with less errors than the human brain. This means that criterion 1 would be perfectly fulfilled. However, considering criterion 2, it becomes clear that the applied logic circuits have only very little to do with how the brain would solve such addition and subtraction problems. Therefore, in the field of Brain-Like AI, such a “simplistic” logic circuit approach would simply not be considered as a good solution and result. Of course, on the other hand, it would not make sense to realize addition and subtraction problems with Brain-Like AI approaches as apparently more “conservative” approaches have proven to even surpass the performance of the human brain in this task. Thus, a selection of the right challenges for which Brain-Like AI could really bring significant benefits is essential.

4. Examples for Brain-Like AI Architectures and their Application

4.1 Objective and Focus of Presented Developments

As outlined in Chapter 2, Brain-Like Artificial Intelligence could be a way to introduce a paradigm shift in certain fields of automation and speed up progress in the development of machines with human-level perceptual and cognitive skills. Currently, we are in the initial phase of this process. The first and foremost objectives of this first phase, which might even constitute the most significant challenges of the whole path, are:

- To actually manage to derive sound and technically implementable models concerning different brain functions from the currently still incomplete and partly contradicting body of knowledge provided by different disciplines of brain science.
- To demonstrate that applying such brain-inspired architectures for automation is principally feasible and can be beneficial.

Accordingly, this chapter concentrates on those two objectives. A quantitative comparison of the performance of the developed models with other possible technical approaches for different

[1] Eventually, a comparison to other technical systems could be done additionally if technical systems are available that can approach the given task.

[2] It is important to note that this does not necessarily mean that the similarity of the developed model and the brain need to be taken down to the lowest neural and chemical processing levels. On the contrary, keeping a certain level of abstraction in order to not be mired in unnecessary details, which might merely derive from the fact that the brain is based on a biological substrate, is in many cases advantageous.

application domains is currently out of the scope. Such a comparison is considered of importance only once the more essential challenges and questions have been successfully faced. In the following, three brain-inspired AI Architectures developed by our research group are presented that address the topics of machine perception (Section 4.2), affective situation assessment (Section 4.3), and autonomous decision making (Section 4.4). Section 4.5 furthermore demonstrates how the development of a brain-inspired AI architectures could significantly extend the knowledge in brain sciences and contribute to new hypothesis about brain functioning.

4.2 Human-Like Machine Perception in Complex Environments

This section is concerned with presenting a brain-inspired AI architecture for machine perception. Parts of the underlying model have already been presented and discussed in former work (see for instance [47, 54, 56]). The objective of this section is to concisely summarize and review these insights and discuss them in the context of Brain-Like AI architectures for automation. The architecture can be taken as a "stand-alone solution" for different machine perception tasks in complex environments [53, 55] or it can be integrated into automation systems needing advanced perception/recognition skills like for instance the decision-making model presented in Section 4.4.

4.2.1 Architecture Description

Figure 6 gives an overview about the developed architecture for human-like machine perception which bases on insights about the working mechanisms of the human perceptual system [46]. The central element of the model is the so-called "neuro-symbolic network", which processes data coming from different sensor sources and additionally considers information coming from "higher-level" sources referred to as memory, knowledge, and focus of attention [54]. Within the neuro-symbolic network, so called "neuro-symbolic information processing" takes place based on information exchange of "neuro-symbols". The focus in this article will be on the description of the functioning of neuro-symbols and the neuro-symbolic network. Details about the other modules and functional aspects of the model can amongst others be found in [30, 46, 49, 51, 60].

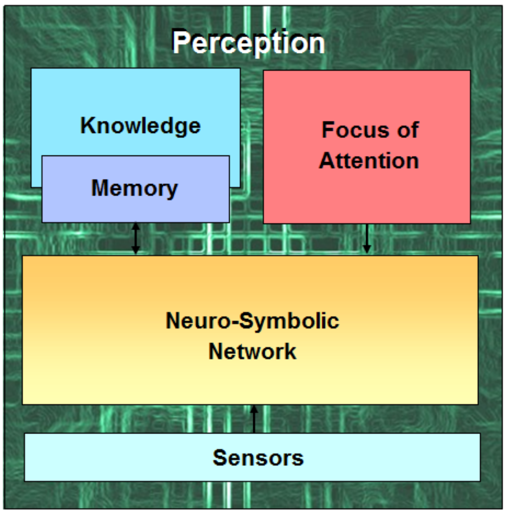
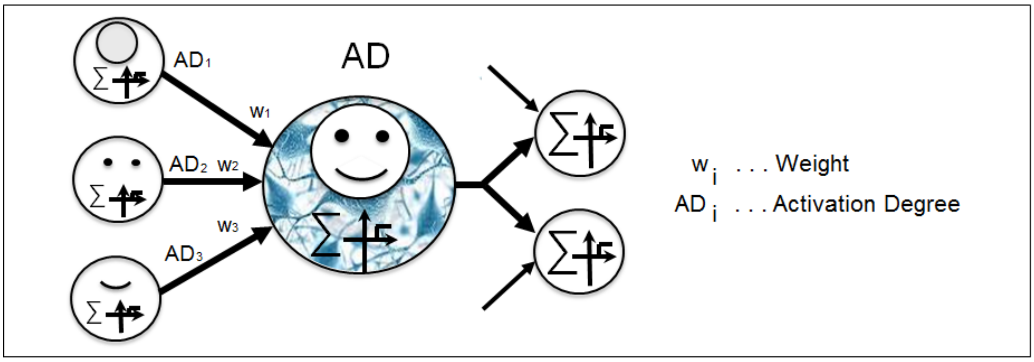


Figure 6. Overview of Brain-Inspired Architecture for Machine Perception

4.2.1.1 Function Principle of Neuro-Symbols

In the neuro-symbolic network, neuro-symbols act as basic information processing units. The idea of using neuro-symbols as basic processing elements came from the following observation: In the brain, information is processed by interconnected neurons and their specific firing patterns. Nevertheless, in the mind, processing of information is considered as a symbolic process based on symbols and their manipulation [19]. Accordingly, neural and symbolic information processing in the brain/mind can be considered as belonging to one and the same phenomenon, however, on two different levels of abstraction.

The first question that was investigated in our work was what connection or interface exists between the described neural and symbolic level. This investigation was made in the context of human perception. In the case of perception, “symbols” refer to perceptual images like faces, persons, specific sounds, melodies, voices, textures, shapes of objects, particular smells and tastes, etc. Studying neuroscientific literature, we made a highly interesting observation: It has been reported about neurons in the brain that seem to fire exclusively as a response to specific perceptual images like, e.g., stripes of a certain orientation and length, sounds of specific frequencies, faces, specific melodies, etc. [20, 23, 27]. This observation led us to the development of neuro-symbols as basic information processing units of our brain-inspired perception architecture.



$$AD = \frac{1}{\sum_{i=1}^{n} wi} \cdot \sum_{i=1}^{n} wi \cdot ADi$$

```
if (AD > Threshold)
        Activated Neuro-Symbol
        Transmit AD
else
        Deactivate Neuro-Symbol
        Transmit 'zero'
end
```

Figure 7. Function Principle of Neuro-Symbols

In Figure 7, the basic function principle of neuro-symbols is illustrated. One characteristic of neuro-symbols is that they represent symbolic information. In the case of perception, this symbolic inforamtion are perceptual images like for instance a face or a voice (see Section 4.2.1.2 for more details). Furthermore, neuro-symbols show a number of analogies to biological neurons. They have an activation degree (AD), which indicates if the perceptual image that each neuro-symbol respresents is currently perceived in the environment. Each neuro-symbol has a certain number of inputs and one output. Via the inputs, information about the activation degree of other neuro-symbols is collected. Like illustrated in the example of Figure 7, a neuro-symbol representing a face could for instance receive information from neuro-symbols representing a head, eyes, and a mouth.

The information of the inputs can be weighted according to the reliability of the information from different input sources. This can for example be useful if the information comes from different types of sensory receptors (visual, acoustic, tactile, etc.). Furthermore, weights can be negative to realize inhibitory functions (see Section 4.2.1.2). As illustrated in Figure 7, the activation degree of a neuro-symbol is calculated by suming up the weighted activation degrees (wi • ADi) of all inputs and normalizing this value to the sum of the weights of all inputs. If the calculated activation degree exceeds a certain threshold value, the neuro-symbol is activated meaning that the perceptual image it represents has been detected in the environment. The information about the activation degree of the neuro-symbol is then transferred via the output to other neuro-symbols to which it is connected.

4.2.1.2 Neuro-Symbolic Networks

In order to perform complex tasks, neuro-symbols have to be connected to neuro-symbolic networks. For the structural organization of this neuro-symbolic network, the modular hierarchical organization of the human perceptual cortex as described by A. Luria [27] was taken as a blueprint (see Figure 8).

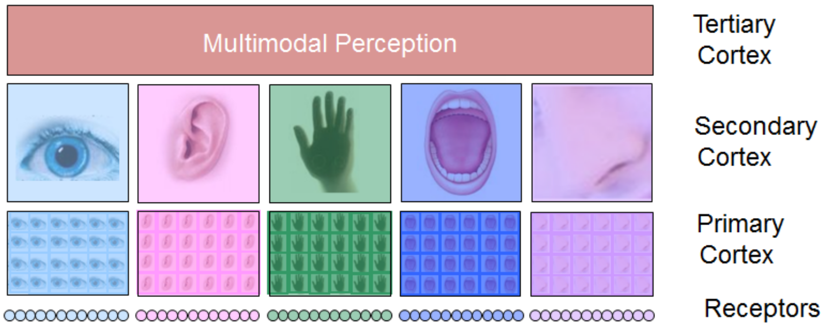


Figure 8. Modular Hierarchical Organization of the Human Perceptual System

According to A. Luria [27], the starting point for perception is information from different sensory receptors (visual perception, acoustic perception, somatosenrory perception, taste, and smell) which is then processed in at minimum three processing stages corresponding to the primary, secondary, and tertiary cortex, which can be composed of several sub-layers. In the primary and secondary cortex, information of each sensor modality is processed separately and in parallel. In the primary cortices, relatively simple information is extracted from the sensory receptors. For the visual modality, this could for example be lines of a certain orientation or lines moving into a certain direction with a certain velocity. For the acoustic modality, this could be a sound of a certain frequency. In higher layers, these features are processed to more and more complex perceptual images which then result in the secondary cortex in complex unimodal perceptions. For the visual modality, this could for instance be a face or a person. In the acoustic modality, this could be the perception of a voice. In the tertiary cortex, information of all modalities is merged to a multimodal perception of the environment. An example would be to combine information form the visual and acoustic modality (e.g., a face and a voice) to conclude that a particular person is currently talking.

In analogy to the modular hierarchical organization of the perceptual system of the brain, neuro-symbols are structured to neuro-symbolic networks as depicted in Figure 9. In the network, incoming sensory information is processed in different layers to more and more complex "symbolic" information until they result in a multimodal perception of the environment. The feature level corresponds to the processing in the primary cortex. The sub-unimodal and the unimodal level correspond to the processing in the secondary cortex. Finally, the multimodal level and the scenario level emulate the processing of the tertiary cortex. Details concerning the function of the different levels have been described in [47, 53, 55]. Neuro-symbols of a lower level always constitute the "neuro-symbol alphabet" of the next higher level. Different combinations of these lower level neuro-symbols result in an activation of different higher-level neuro-symbols. In addition to these feedforward connections, neuro-symbolic networks can have feedback connections and can receive input from knowledge and memory and focus of attention (see the example of Figure 11).

Concerning the sensor modalities, there can be used sensor types that have a correspondence in the human sense organs (e.g., video cameras for vision, microphones for acoustic perception, light barriers, temperature sensors, motion detectors, or contact sensors for somatosensory perception, chemical sensors for olfactory perception) or there can be used sensors that have no analogy with human sensing modalities (e.g., sensors measuring electricity or magnetic fields).

In analogy to how it is reported for the brain by A. Luria [27], connections of the lowest levels of the architecture of Figure 9 are predefined. Higher-level connections are set via a learning process, concretely via a supervised learning process that was described in detail in [46]. More recent research findings indicate that learning could also already take place at lower levels of perception and that unsupervised learning could be crucial for setting these connections. In [51], first attempts have been made to develop an unsupervised learning strategy for the model.

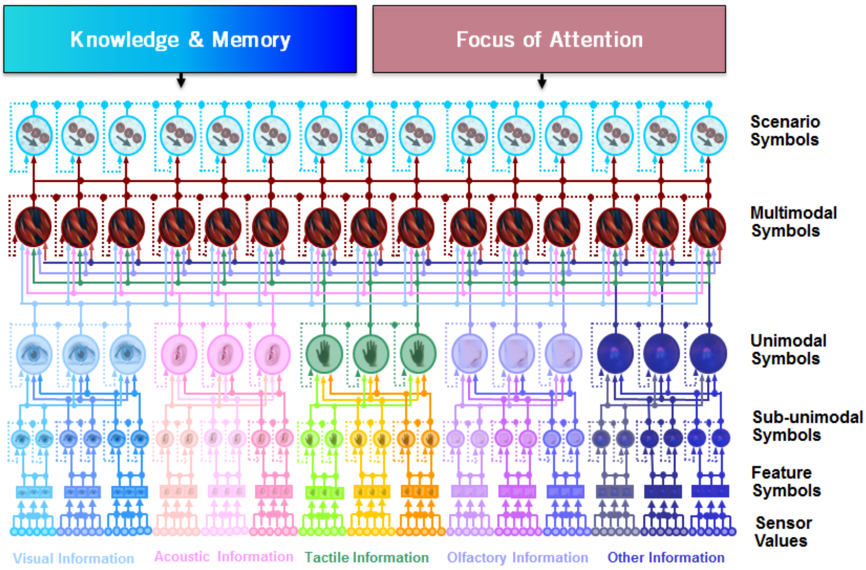


Figure 9. Modular Hierarchical Organization of Perceptual Neuro-Symbolic Networks

To illustrate the basic working principle of a perceptual neuro-symbolic network in a concrete application, a simplified, concrete example is given in the following. In this example, different activities of persons in a room shall be detected. For this purpose, an office meeting room is equipped with different sensors (tactile floor sensors, motion detectors, light barriers, a door contact sensor, a camera, and a microphone) as sketched in Figure 10.

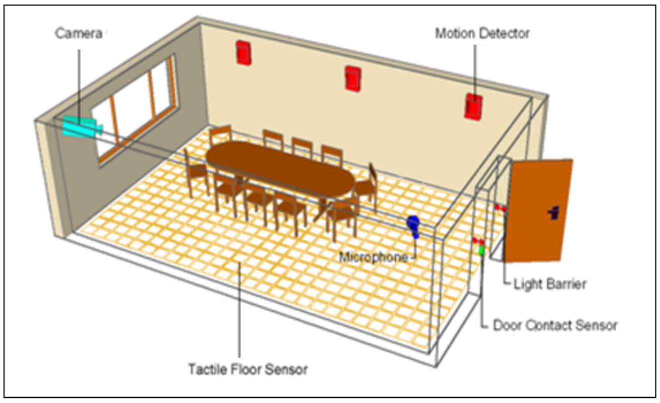


Figure 10. Meeting Room Equipped with Different Sensors

Depending on the different activities that shall be detected in the meeting room, a neuro-symbolic network is configured based on specific training data (see for instance [53, 55, 59]). Figure 11 shows the neuro-symbols of a neuro-symbolic network after training that are employed for detecting that a person walks around inside the room. If a person walks around inside the room, the motion detectors detect “motion”. Furthermore, the tactile floor sensors can detect an object, which is represented by the neuro-symbol “object present”. The information of these two neuro-symbols is then combined to the neuro-symbol “object moves”.

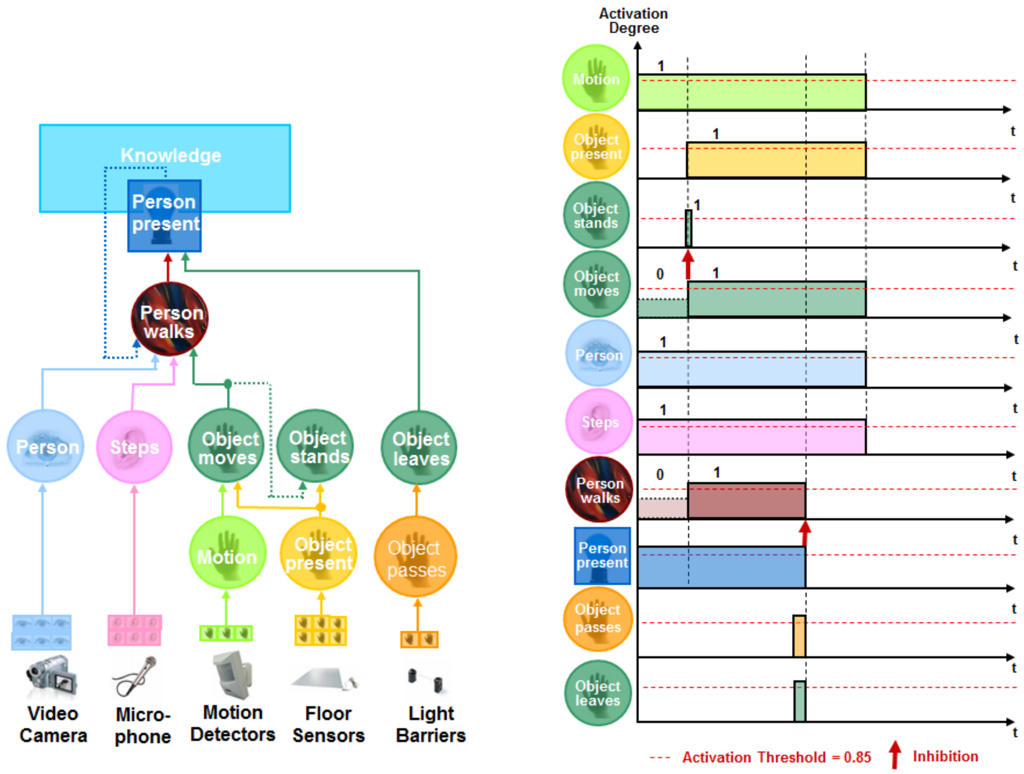


Figure 11. Activated Neuro-Symbols for Detecting that a Person walks around in the Room

With the example of Figure 11, also the function of feedback connections can be explained. According to the existing feedforward connections, the neuro-symbol “object stands” would be activated together with the neuro-symbol “object moves” whenever the neuro-symbols “motion” and “object moves” are active, because it is activated by a subset of the neuro-symbols that activate the neuro-symbol “object moves”. This activation would however be undesired in this concrete case. For this reason, an inhibitory feedback connection exists from the neuro-symbol “object moves” to the neuro-symbol “object stands” that inhibits the activation of the neuro-symbol “object stands”.

In addition to the “tactile modalities” just mentioned, the visual image of a “person” is detected via the video camera and “steps” can be perceived via the microphone. On the next higher level, this leads to the activation of the neuro-symbol “person walks”.

With the example of Figure 11, it can furthermore be shown how the interaction between the neuro-symbolic network and so-called “memory symbols” and “knowledge” looks like. For this purpose, it is assumed that the person walks out of the room (leaves the room) but is still in the near vicinity of the door. In this case, the person is still inside the detection range of the video camera, the microphone, the motion detectors and the tactile floor sensors, because they also cover a small area outside the room in the vicinity of the door. Therefore, based on current sensor information only, the network would still detect a person walking around in the room, which would however be incorrect. Such an incorrect activation of the neuro-symbol “person walks” can be circumvented by information from a pair of light barriers together with “memory symbols” and “knowledge”. Memory symbols can store information about events that happened in the past that are relevant for the current of future situation. In the example of Figure 11, the light barriers are triggered if the person leaves the room leading to a short activation of the neuro-symbol “object passes” and as a consequence the neuro-symbol “object leaves”. The neuro-symbol “object leaves” is connected to the memory symbol “person present”. This memory symbol is activated as soon as a person enters the room and is deactivated when the person leaves the room again. The memory symbol “person present” is then connected to the knowledge module, which can reason that if no person is present in the room any more, no person can walk around in the room. This way, an inhibitory connection to the neuro-symbol “person walks” deactivates this neuro-symbol after the person has left the room.

Of course, if it shall only be detected that a person walks around in the room like in the demo example from Figure 11, it would not be necessary to use so many different sensor types showing a high degree of redundancy. However, if not just one but a broad range of activities in the room or in a whole building shall be recognized, redundancy and fault tolerance in information is necessary to avoid false perceptions [46].

In addition to the functions described in this article so far, neuro-symbols can have so-called properties, which specify the perceptual images they represent in more detail. An important example for a property would be the location where a complex perceptual image has been perceived. The functionality of properties will not further be handled in this review. Details can be found in [47]. Furthermore, neuro-symbols cannot only process input information reaching at the same instant of time, but can process information reaching in a certain time window or in a certain temporal succession. Details about temporal aspects of neuro-symbolic information processing have been described in [30]. Details about the mechanism of focus of attention, not furhter mentioned in this article, can be found in [47, 60]

.

4.2.1.3 Neuro-Symbolic Networks versus Neural Networks

After having briefly illustrated the basic function principle of neuro-symbolic networks, this section aims at reviewing their affinities and differences to standard neural networks like for example multi-layer perceptrons (MLPs) [58]. A summary of these affinities and differences is given in Figure 12. The affinities concern certain functions of individual nodes of the networks. In both cases, weighted input information is summed up and an activation function is applied to this sum. In both cases, the individual nodes are interconnected to form networks. Much larger than the number of affinities between neuro-symbolic networks and neural network is however the number of differences. The first difference consists in the application domain. Neuro-symbolic networks have so far mainly been applied for complex, large-scale sensor data processing of multimodal data – an application which can so far barely be handled by neural networks [38]. In neuro-symbolic networks, each node of the network has a particular symbolic meaning which makes comprehension of the internal processes easy. In contrast, neural networks have been described as black box models concerning the meaning of their nodes and weighted connections and thus it is generally just practicable to observe their input-output behavior. A further distinction criterion is that neuro-symbolic network can have so-called properties, which specify the corresponding neuro-symbols in more detail. An example for a property would be the location where a perceptual image had been perceived. Details about properties and their function and benefits have been described in [46]. Additionally, neuro-symbolic networks cannot only process information being available at a particular instant of time but can process information that arrives within a certain time window and information that arrives in a particular sequence. These aspects have been described in [30]. Different from most neural networks, where usually each node of one layer is initially connected to each node of the next layer, neuro-symbolic networks show a modular hierarchical organization that is inspired by the structural organization of the human perceptual cortex [20, 27]. The modular hierarchical organization of the neuro-symbolic networks furthermore allow the development of hybrid concepts, i.e. the combination of the brain-inspired concepts with more classical processing principles like e.g., the substitution of the visual modality with classical pattern recognition algorithms.

In neural networks, the function of the weights of the pre-connected nodes is to map input values to desired output values during a learning process. In contrast, in neuro-symbolic networks, weights have the function to represent different reliabilities of information sources (e.g., if one sensor modality is more reliable than another one). Furthermore, as explained above, neuro-symbolic networks can have inhibitory feedback connections between nodes which have the function to suppress undesired activations of neuro-symbols and they allow the integration of concepts like memory, knowledge, and focus of attention [60].

Differences also exist in the learning mechanisms of neural and neuro-symbolic networks. Instead of an alteration of the weights, learning in neuro-symbolic networks concerns the setting of connections, the determination of correct temporal successions of input signals that shall lead to an activation, and the determination of value ranges for properties that shall lead to an activation in a particular situation (see [46] for more details). Additionally, it is possible to remove redundant elements from the network during the learning phase or to add elements in case a more detailed distinction between neuro-symbolic nodes is necessary. Further details about the learning mechanisms of neuro-symbolic networks can be found in [50, 51].

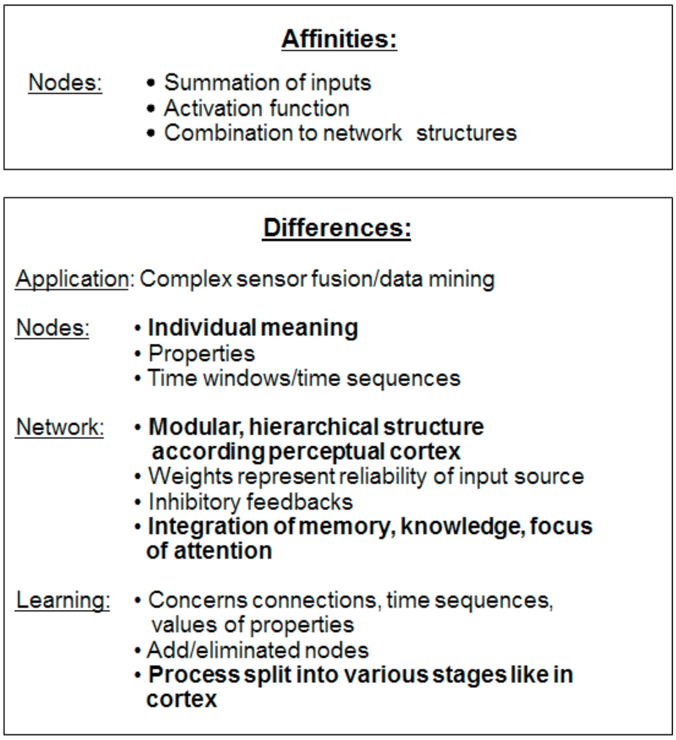
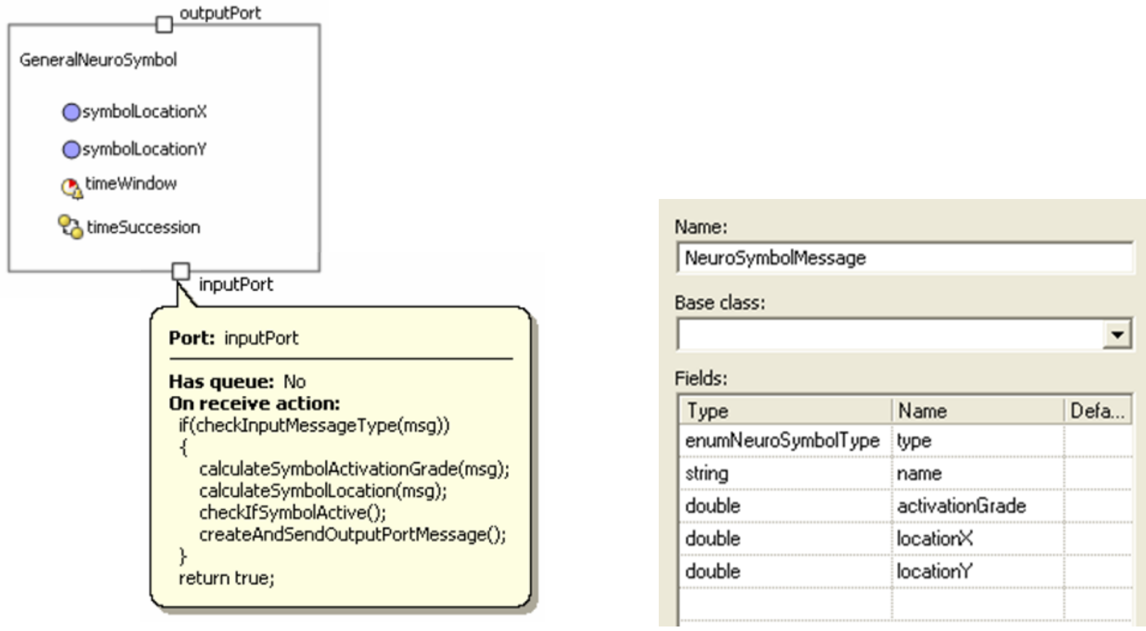


Figure 12. Affinities and Differences of Neuro-Symbolic Networks in Comparison to Classical Neural Networks (e.g., MLPs)

4.2.2 Implementation and Application

For validating the function of the architecture presented in Section 4.2.1, it was implemented in two different programming languages. These two implementations were independent from each other. On the one hand, AnyLogic [1], a graphical fast prototyping tool based on Java was used. On the other hand, a programming language based on RuleML [4] was applied. Details of these implementations have been presented in [47, 53]. For showing the basic implementation concept, here, a short overview about the implementation in AnyLogic is given followed by references to practical test applications.

Figure 13 and Figure 14 show screenshots of the model implementation in AnyLogic. Figure 13a shows how individual neuro-symbols were implemented. Neuro-symbols are realized by so-called active objects with an input port and an output port via which information is exchanged with other elements. Additionally, variables are used for calculating the activation of the neuro-symbols (not depicted) and for storing properties of neuro-symbols (e.g., the location property). Timers and state charts serve for processing information that arrives in a certain time window or in a certain temporal succession at the input port. Whenever new input information arrives at the input port, the activation degree of the neuro-symbol is recalculated and checked against the threshold value. Based on this, the neuro-symbol is either activated or deactivated and the corresponding information is sent via the output port by using "message objects" (see Figure 13b). These message objects again contain different variables, amongst others to store information about the activation of neuro-symbols and the value of their properties.



(a) Active Objects for Neuro-Symbol (b) Message Objects for Communication between Neuro-Symbols

Figure 13. Implementation of Neuro-Symbols and their Communication in AnyLogic

To perform complex functions, individual neuro-symbols are then connected to networks. Figure 14 shows a screenshot of the AnyLogic implementation of the overall system at the beginning of the learning phase. The lowest neuro-symbolic levels receive the direct sensor information as input. The higher neuro-symbolic levels are originally not interconnected amongst each other. Instead, they are connected to so-called “learning ports”, which additionally receive control information needed for the supervised learning process. Details about the multi-stage multi-level learning process can be found in [47].

In the following, a brief summary of a number of successful first test applications of the described perception architecture is given. In [50], the architecture was applied for a scenario recognition system to monitor the activity of persons in a building using video data, audio data, and information from different “tactile” sensors. In [59, 62], the topic of energy management in building automation systems and prosumer buildings was targeted. For this purpose, a model combining brain-inspired recognition mechanisms with rule-based decision-making concepts for actuator control was implemented and tested. For this purpose, activities of occupants and scenarios going on in the building were detected, which served as input for the decision-making unit for determining the appropriate actuator control. The task that was fulfilled by these systems was the control and reduction of the energy consumption (lighting, heating, air conditioning, etc.) in an office building without affecting the occupants' comfort and to efficiently manage renewable energy resources. In [4], again, the domain of building automation was targeted, this time for the development of an activity-dependent alerting system. For this purpose, the brain-inspired perception/recognition model was combined with a rule-based decision-making system and implemented in RuleML. The task of this alerting system was to detect ongoing activities, events, and situations in an office building equipped with sensors and to decide which of these activities to communicate to particular occupants via different electronic media (PC, mobile devices, screens, etc.) to inform them about relevant opportunities, safety and security relevant issues, to encourage them in energy and resource saving behavior, and so on. The feasibility of employing the brain-inspired model for these tasks was demonstrated via both a hardware-based test bed (an office kitchen equipped with sensors) and simulations (a whole virtual office floor equipped with virtual sensors). Finally, our latest work in this field is concerned with employing the introduced model for human activity recognition in elderly homes and comparing the achieved recognition results with available classical AI approaches [63].

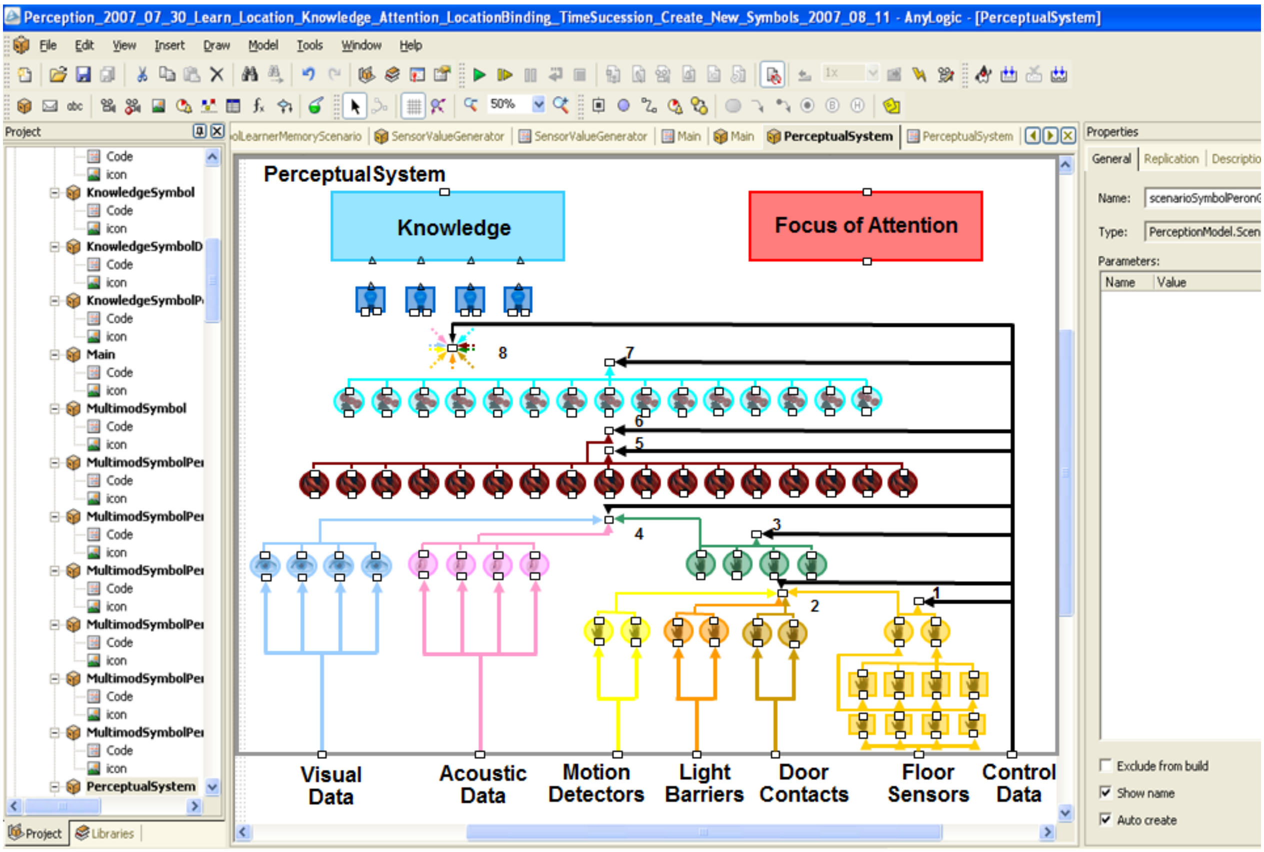


Figure 14. Implementation of Overall Architecture of the Perception Model in AnyLogic

4.3 Affective Situation Assessment

This section is concerned with extending the brain-inspired perception architecture presented in Section 4.2 by brain-inspired situation assessment mechanisms using a concept called emotions. The field of affective computation in AI (i.e. the development of computational models including the concept of emotions) started to develop after brain researchers had noticed that emotions play a crucial role in intelligent behavior and are involved in cognitive problem solving and decision-making [11, 29]. Thus, an increasing number of AI researchers have started to believe that computational models of emotions would also be needed to design intelligent systems. One of the most famous statements in correlation to this is the one of the AI pioneer M. Minsky [36]: *"The question is not whether intelligent machines can have emotions, but whether machines can be intelligent without any emotions."* A. Sloman [40] points out that *"the need to cope with a changing and partly unpredictable world makes it very likely that any intelligent system with multiple motives and limited powers will have emotions."*

In computational models, emotions are employed as a value system for giving processed concepts meaning in terms of a classification as good or bad or with a finer breakdown. However, as outlined in [26], modeling emotions in technical terms has turned out to pose many difficulties and has often been deemed as just not feasible. In [52], we analyzed the reasons for these implementation difficulties not based on "abstract" theoretical discussions but on a concrete implementation attempt. In the following, a summary of the insights gained during this development and analysis process is given.

4.3.1 Architecture Description

As indicated above, the model of affective situation assessment constitutes an extension of the neuro-symbolic perception model presented in Section 4.2 by integrating the concept of emotions as evaluation mechanisms. With this, it becomes possible to assign "meaning" to perceptual images and to provide an interface for decision-making and action execution processes (see Section 4.4). Before describing details about the developed architecture, it shall at this point first be briefly explained how emotions can be understood from a computational point of view. Further details concerning this topic can be found in [10, 52].

In brain science literature, not just one definition for emotions can be found but various. In principle, emotions can be regarded as an evaluation mechanism that classifies information as good or bad or tags it with finer range of valuations (beneficial, pleasant, precarious, dangerous, etc.). The advantage of evaluating information using emotions in contrast to using "logical" reasoning processes seems to be that emotional evaluations can be achieved very fast. As a consequence, certain decisions can be taken faster without time consuming explicit analyses of the situations and all its possible outcomes, implications, and consequences. Furthermore, an emotion can even directly result in the triggering of an action (e.g., to prevent an individual/system from harm).

In addition, [29, 41] describe emotions as an internally directed perception process, implying that emotions can be regarded as a further, internally directed sensor modality providing information about the current state of the "bodily self". Thus, while other sensor modalities are concerned with acquiring information from the environment, emotions provide information about processes going on in the body of an individual. [11] indicates that apart from body states, emotions can be triggered from objects or events perceived in the environment and from higher cognitive processes, e.g., if the individual realizes that a catastrophe is about to happen. Furthermore, it seems plausible that certain emotions can facilitate the activation of other emotions.

[29, 41] distinguish between so called basic/primary emotions and complex/secondary emotions in humans and animals. They describe basic/primary emotions (e.g., rage, fear, panic, and seeking) as being hardwired, meaning that an individual responds with such an emotion in a pre-organized fashion when particular key features (e.g., size, large span, type of motion, certain sounds, certain configurations of body states) or their combination are perceived in the environment or the body. These emotions can then trigger body responses, without the individual necessarily being aware of these responses. These responses increase the likelihood of the organism to survive in the environment. Complex emotions are described as being learned emotions. Via complex/secondary emotions, it becomes possible to set an association between an emotion and the object that triggered it (e.g., feeling upset because the thing over there bit me; feeling hungry and wanting to eat that thing over there).

Based on the descriptions given above and the concept of neuro-symbolic information processing outlined in Section 4.2, so-called "affective neuro-symbols" were defined for the affective situation assessment architecture (see Figure 15). These affective neuro-symbols can principally receive information from four different sources: (1) body states, (2) objects and events perceived in the environment (external perception), (3) from other emotions and (4) cognitive (reasoning) processes. An input from one of these sources can in certain circumstances already be sufficient to activate an affective neuro-symbol. Different sources can either have an exhibitory or inhibitory effect on the activation of an affective neuro-symbol.

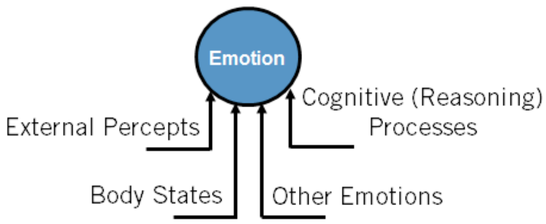


Figure 15. Input Sources of an Affective Neuro-Symbol Representing an Emotion

Based on the concept of affective neuro-symbols, a model was developed according to which emotions can be represented by affective neuro-symbolic networks (see right half of Figure 16, referred to as architecture of "internal perception" in contrast to the "external perception" architecture of the left half of Figure 16, which has already been presented in Section 4.2).

The individual affective neuro-symbols (depicted as circles) represent different emotions (fear, anger, guilt, joy, rage, panic, love, happiness, etc.).

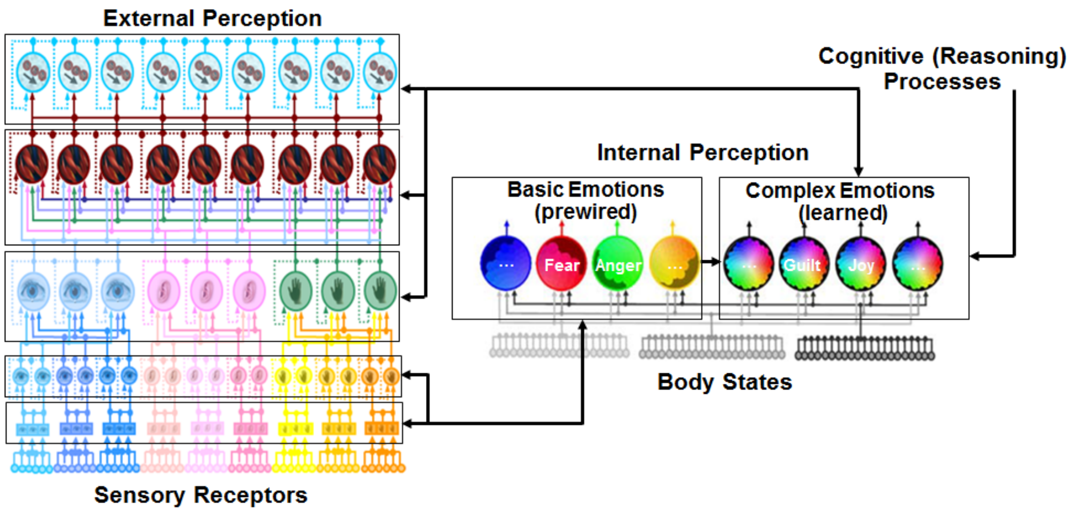


Figure 16. Architecture for Affective Situation Assessment of Perceptual Images (Internal Connections between Emotions are not Depicted for Better Clarity of the Graphic)

As already indicated in Figure 15, emotions can be triggered from different input sources (from the body, from external perception, from cognitive processes, and from other emotions). Activated emotions can then activate other systems like for instance the motor cortex for triggering certain movements and behaviors. Some details concerning this topic are described in the decision-making model in Section 4.4.

In the model, it is distinguished between basic (primary) and complex (secondary) emotions. One question that has been analyzed was if – similar as for the model of external perception presented in Section 4.2 – affective neuro-symbols can also be arranged in a hierarchical fashion (as the names primary and secondary emotions might imply). In this case, several primary emotions would always serve as input for secondary emotions. However, a more detailed analysis of affective neuroscience literature indicated that in the developed brain, these two types of emotions rather seem to co-exist one next to the other (although primary emotions seem to be necessary for the evolvement of secondary emotions).

As indicated above, basic emotions are "prewired emotions" and complex emotions are "learned emotions". Therefore, apart from body states, basic emotions are connected to prewired parts of the architecture of external perception, i.e. the lowest neuro-symbolic levels of this neuro-symbolic network and already exist at system startup (equaling the time birth of an organism). In contrast to this, complex emotions are connected to the higher levels of the neuro-symbolic network of external perception, which only develop their connections during learning processes. More details concerning this learning process can be viewed in [52]. Furthermore, complex emotions can be triggered from concepts established in higher cognitive levels outside perception.

4.3.2 Implementattoon and Application

With the model presented in Section 4.3.1, a computational model for affective situation assessment has been introduced conclusively combining low-level (bodily) und high-level (cognitive) concepts of emotions and illustration changes of emotional mechanisms over development. The model was implemented in AnyLogic and analyses were made concerning its potential for application in both situation-aware building automation systems (see Section 4.2) and autonomous agents (see Section 4.4). Details about these topics can be found in [52].

By actually targeting an implementation of such a brain-inspired affective computational model for different applications, it was identified that, in addition to focusing on the cognitive aspects of emotions, it is crucial to consider the bodily aspects of emotions and their grounding in a visceral body. This is particularly important if a correlation between emotions and perceptual images or cognitive concepts are not pre-defined but have to be acquired through a learning process. In the brain, a learning of the correlations between emotions and perceptual images is not possible without bodily grounding. As a consequence, when aiming at simulating and emulating the human (or animal) brain, efforts will most likely have to be spent on simulating and emulating the functions of the corresponding body to a certain degree. This insight presented in detail in [52] has already been taken up by T. Deutsch [14]. In the end, the controversial assertion that the mind and the body are a single whole seems once more be proven to be certain when going beyond more "basic" computational implementations of human behavior.

4.4 Autonomous Decision-Making in Complex Environments

This section is concerned with presenting a brain-inspired AI architecture for autonomous decision-making. Parts of the underlying model have already been presented and discussed in former work (see for instance [28, 34, 56, 59]). The objective of this article is to concisely summarize and review these insights and discuss them in the context of Brain-Like AI architectures for automation.

4.4.1 Architecture Description

An overview of the decision-making architecture is presented in Figure 17. The architecture was guided by two core concepts. The first core concept is that human intelligence bases on a combination of low-level and high-level mechanisms. Low-level mechanisms are mainly predefined. They are not in all situations completely accurate but have the advantage of being fast. High-level mechanisms are not predefined and thus slower but more accurate. The second core concept concerns the use of so-called emotions as mechanism for the evaluation of information.

The basic functioning of this architecture is now described in the following step by step using Figure 18a-f, where always the relevant modules of the model are highlighted for better comprehension.

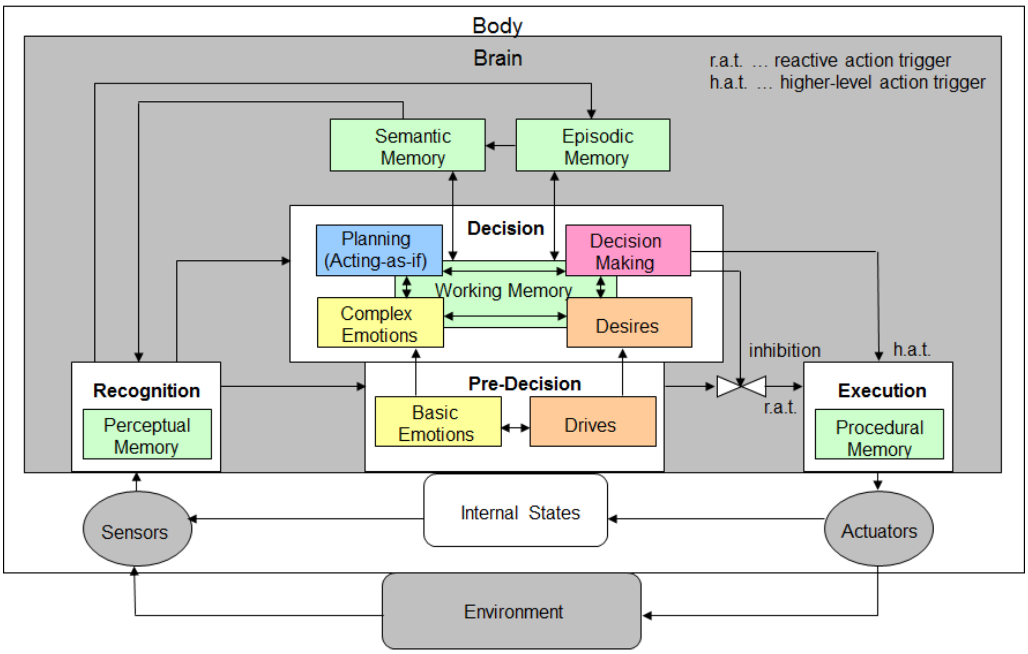


Figure 17. Autonomous Decision-Making Architecture

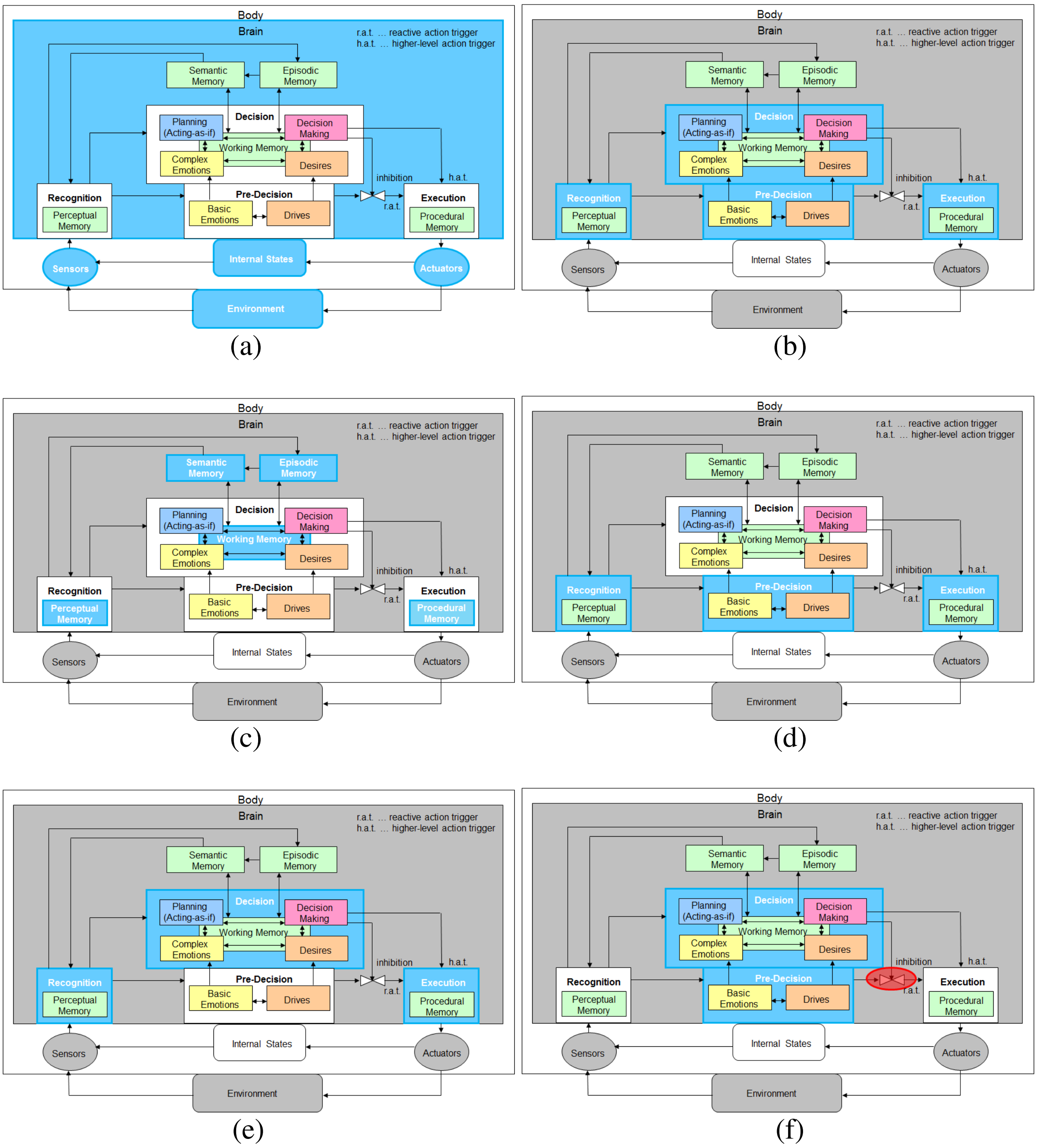


Figure 18. Different Modules involved in the Autonomous Decision-Making Process

As indicated with the highlighted modules in Figure 18a, the basic schema of the model follows the same principle as the one of Figure 3. Starting point are again particular processes, which can take place in the **environment** of the system or in the system itself (**internal states**). These processes are detected via **sensors** and influenced via **actuators**. The processing of the information takes place in a processing architecture inspired by the human **brain**. The sensors, actuators and brain processing unit are embedded into the physical **body** of the system.

Figure 18b highlights the core function blocks of the architecture. The **recognition** unit is responsible for the processing of sensory information. The pre-decision unit and the decision unit are in charge of taking appropriate decisions based on the perceived current situation. The **pre-decision** unit bases mostly on fast pre-defined lower-level processes that shall keep the system from harm via fast reactions in critical situations. In contrast, the **decision** unit bases on higher-level reasoning processes that need more time but are more flexible and multi-faceted. The **execution** unit is responsible for putting the decisions into action by preparing the control signals for the actuators.

The processes of recognition, decision-making, and execution just described make use of different memory systems, which are highlighted in Figure 18c. The **perceptual memory** stores information like the appearance, emitted sound, form/texture, smell and taste of objects, which is needed to recognize objects and situations (see also Section 4.2). The **procedural memory** contains information about how particular movements have to be carried out. The **working memory** is responsible for temporally storing information that is currently relevant to facilitate a merging with other information and, if appropriate, a transcription to other memory systems. The **semantic memory** stores factual knowledge about the environment and the system itself. The **episodic memory** contains recollections of subjective experiences [13]. It can be considered as the autobiography of the system. Via consolidation processes, episodic memory can in certain circumstances be transformed into semantic memory.

Figure 18d and Figure 18e highlight the main modules involved into a complete decision-making and behavior-selection cycle. First, stimuli from the environment and the "body" of the system serve as input for the recognition unit, where the incoming information is processed. The processed perceptual information (in form of "perceptual images") is then directed to the pre-decision unit. Here, the perceptual images trigger certain basic emotions and drives. **Basic emotions** can be considered as an evaluation mechanism. By emotions, information is classified as good and bad or assessed with a more differentiated evaluation scale (see Section 4.3 for more details).

Drives can be regarded as a prioritization mechanism for possible actions based on the current needs of the system. For example, if the energy level of the system is low, actions and activities related to "energy intake" are ranked high in the priority list. Via basic emotions and drives, the pre-decision can perform a fast selection of certain, mainly pre-defined behaviors, which are then immediately executed via the execution unit. If the pre-decision unit does not trigger any actions based on the currently perceived situation, the information from the recognition unit and the pre-decision unit are transferred to the decision unit. Here, again the decision-making process is influenced and guided by evaluation and prioritization mechanisms, this time by so-called **complex emotions** and **desires**. The concept of complex emotions shows strong similarities to the concept of basic emotions and the concept of desires to the one of drives. The difference is that they are acting in a more complex overall context where for example also social aspects of behavior become of importance. The decision unit furthermore makes use of information from episodic and semantic memory. In the episodic memory, it is searched for situations experienced earlier that show similarities to the current situation and it is analyzed what decisions and behaviors were either beneficial or unfavorable in this context. The semantic memory can additionally support the decision process by providing factual information about what is in general recommend to be done in a certain situation. If no similar situations can be found in the episodic and semantic memory, the **planning** module is activated. Here, different possible (re-)actions to a situation are simulated and their likely outcome is determined. Based on all this information, the **decision-making** module then selects the presumably optimal action.

In general, the pre-decision unit and the decision unit harmonically work together. However, it can also come to situations where a conflict occurs, i.e., the pre-decision unit and the decision unit propose the triggering of contrasting actions. As the pre-decision unit usually reacts faster than the decision unit, the pre-decision unit would always overrule the decision unit in such cases, which can in certain situations however be disadvantageous. Therefore, as indicated in Figure 18f, particular **inhibition** mechanisms exist so that the decision unit can suppress (re-)actions that are initiated by the pre-decision unit. A simple example where such an inhibition mechanism would be activated is the following. Assume that I grasp for an object which turns out to be very hot and therefore activates my pain receptors that transmit this information to my brain. The immediate action triggered by the pre-decision unit would now be to drop the hot object to prevent my fingers from burns. However, it can be the case that the object is very precious and could be destroyed when falling on the floor. In this case, the decision unit can overrule the pre-selected dropping action and instead guide a controlled putting down of the object at an adequate place.

4.2.2 Implementation and Application

A feasibility proof of the principal functionality of the decision-making architecture described above was performed by a computational simulation, which was originally presented in [59] and is briefly summarized here. For this purpose, a virtual environment was designed in which virtual agents, each having implemented an instance of our decision-making architecture as control unit, can perform different activities.

The implementation was performed in AnyLogic – a fast prototyping tool based on Java with a graphical programming interface that supports agent-based modeling. Figure 19 shows a screenshot of the test implementation. The picture in the middle shows the modules and interfaces of the decision-making architecture which were realized by so-called "active objects" and "ports". On the left side, the implemented modules are listed. In the right lower corner of the figure, the agents and the virtual environment are displayed. The environment comprises different "objects" (food sources, obstacles, predators, other agents, etc.). In order to survive, the agents have to access food sources. However, the accessing of food sources bears difficulties and risks. For instance, obstacles have to be overcome to access potential food sources and this activity can in occasions consume more energy than the food source will provide. In other cases, food sources can only be accessed in collaboration with other agents like for example the hunting of bigger animals. Furthermore, agents can risk to be caught by predators while accessing food sources. Therefore, agents have to adequately decide about what activities to perform and what activities not to perform in a particular situation in order to survive and avoid to be harmed.

By the simulations, the behavior of individual agents could be observed in the different occurring situations. By adjusting different "parameters" of the decision-making architecture for different agents, insights could be gained about the influence of different basic and complex emotions, drives, desires, and prior experiences stored in episodic memory on decisions and their outcome. Examples for this were amongst others presented in [59].

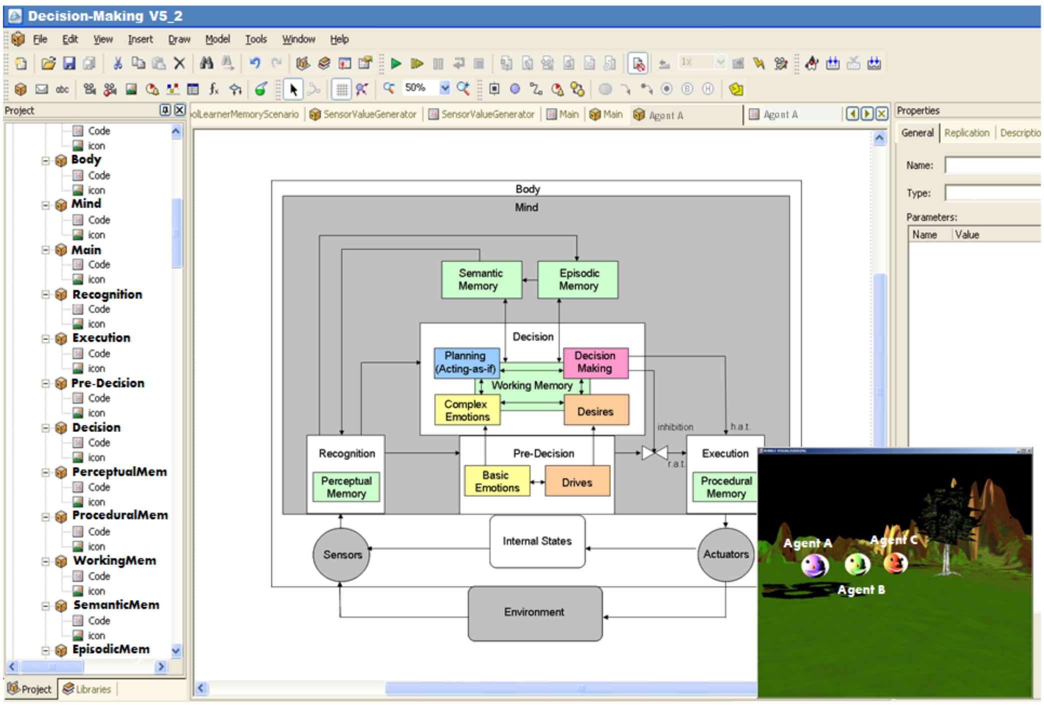


Figure 19. AnyLogic Implementation of Decision-Making Architechture in an Virtual Autonomous Agent Environment

Basing on the insights gained in the autonomous agent application, our most recent work is concerned with applying a slightly adapted version of the decision-making architecture for the purpose of energy management and trading in prosumer buildings and neighbourhoods. The performance of the architecture is compared with a classical AI approach employed for the same purpose [63].

4.5 Example for Implications on Brain Science – A Conclusive Potential Solution to the Binding Problem

By deriving concepts about the structures, information processing principles, and functioning of the brain for the design of more efficient, flexible, and "intelligent" technical systems, not only the knowledge in the field of engineering and computer science was extended in several aspects but also in the domain of brain sciences. In this context, the huge potential of technically implementing such brain-inspired models lies in the fact that in order to achieve an actually functioning technical system, dark spots and inconsistencies in brain theories cannot just be left out and ignored. They have to be clearly identified and if possible tried to be filled. In the following, an example for new insights and hypotheses concerning brain functioning gained during the course of the work presented in this article is outlined. A description of further impacts on the field of brain sciences can amongst others be found in [49, 51, 52].

Based on the developed perceptual model presented in Section 4.2, a conclusive potential solution to the so-called binding problem in perception has been suggested in [60]. The binding problem in general is concerned with finding an explanation for how information in the brain, which is processed in parallel in widely distributed systems, can result in a unified experience. In perception, the binding problem is concerned with explaining how the information coming from millions of sensory receptors – being in a first instance processed separately and in parallel – can in the end result in a unified and unambiguous perception of the world. The binding problem has puzzled researchers in brain sciences for decades and is considered to be one of the key questions to brain understanding [21]. As the perceptual model presented in Section 4.2 emulates the distributed structure of the perceptual system of the brain, this binding problem had to be faced during its development. Various solutions to the binding problem have been suggested in literature so far (see for instance [3, 17, 18, 22, 23, 33, 44, 66]). However, each of them has certain weak points. As a result of our research, in [60], a solution to the binding problem for perception was suggested by combining the already existing binding hypotheses in a conclusive way, supplementing them with other insights about the perceptual system of the brain, and translating them into a technically implementable model. It was demonstrated via computational simulations that different binding mechanisms proposed in literature are not mutually inclusive. On the contrary! At different hierarchical levels and in different development stages, different binding mechanisms are acting in perception. An overview about these circumstances is given in Figure 20. A detailed description can be found in [60].

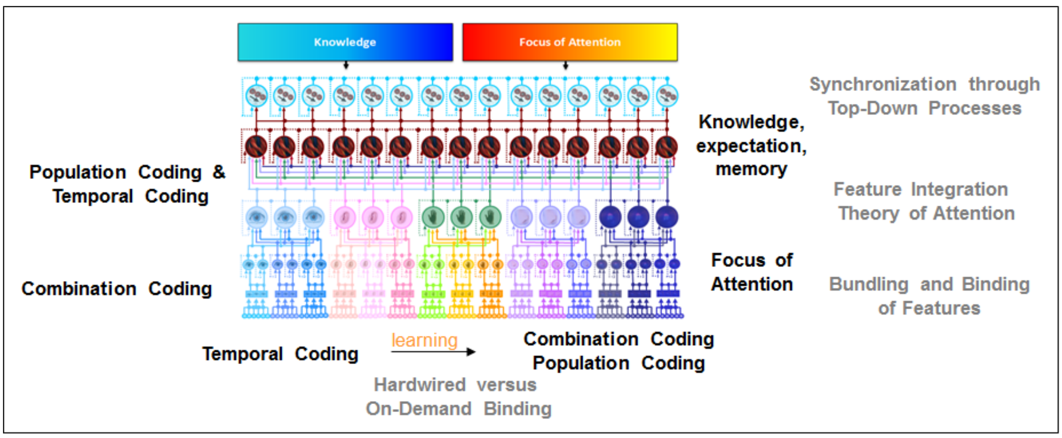


Figure 20. Overview about which Binding Mechanisms Work at what Hierarchical Levels and Development Stages of the Brain

5. Conclusion and Outlook

In this article, an introduction to the field of Brain-Like Artificial Intelligence for automation was given and the principal feasibility of deriving models for complex, robust perception, situation assessment, and decision-making from function principles of the brain was demonstrated. These models were based on concepts like neuro-symbolic cognition, different types of memory, semantic knowledge, focus of attention, body states, emotions, drives, desires, learning, bottom-up and top-down processes, perceptual binding, feedback loops, and inhibition circuits. A feasibility proof of the models was provided by applying them to different applications in the field of advanced, situation-aware building automation systems and autonomous agents. Furthermore, the developed brain-inspired models contributed to an extension of the knowledge in brain sciences, for instance by providing a novel conclusive potential solution to the binding problem, which is a key question in brain research.

These first results indicate that applying Brain-Like Artificial Intelligence for automation is a highly promising domain with great potential. Nevertheless, the field is currently still standing at its very beginning. Many secrets about brain functioning remain to be unveiled and new grounds concerning scientific methodologies have to be broken. To achieve this, engineers will have to join forces with brain scientists and life scientists and carry out research in tight collaboration. Researchers of different domains will certainly need to learn to understand the language of each other up to a certain extent and to think out of the box.

Acknowledgement

The Competence Centre CTR is funded within the R&D Program COMET - Competence Centers for Excellent Technologies by the Federal Ministries of Transport, Innovation and Technology (BMVIT), of Economics and Labour (BMWA) and it is managed on their behalf by the Austrian Research Promotion Agency (FFG). The Austrian provinces (Carinthia and Styria) provide additional funding.

References

[1] AnyLogic (2004). Anylogic User’s Manual. Technologies Company Ltd.

[2] Automation (2011). Automation – History, types of automation, the role of computers in automation, applications, the human impact of automation – in the home. http://science.jrank.org/pages/679/Automation.html, 2011.

[3] Barlow, H.B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology. Perception, 1:371-394, 1972.

[4] Boley, H. (2007). Are your rules online? Four web rule essentials. In Proc. Int. Symp. Adv. Rule Interchange Appl. RuleML, pp. 7-24.

[5] Bruckner, D. (2007). Probabilistic Models in Building Automation: Recognizing Scenarios with Statistical Methods. PhD thesis, Vienna University of Technology, 2007.

[6] Bruckner, D.; Picus, C.; Velik, R.; Herzner, W.; and Zucker G. (2009). High-level hierarchical semantic processing framework for smart sensor networks. In Human-Computer System Interaction: Backgrounds and Applications, pp. 347-358. Springer Berlin/Heidelberg.

[7] Bruckner, D.; and Velik, R. (2010). Behavior learning in dwelling environments with hidden Markov models. IEEE Transactions on Industrial Electronics, 57(11):3653-3660.

[8] Bruckner, D.; Picus, C.; Velik, R.; Herzner, W.; and Zucker G. (2012). Hierarchical semantic processing architecture for smart sensors in surveillance networks. IEEE Transactions on Industrial Informatics, 8(2):291-301.

[9] Burgstaller, W. (2007). Interpretation of Situations in Buildings. PhD thesis, Vienna University of Technology.

[10] Burgstaller, W.; Lang, R.; Poerscht, P.; and Velik, R.(2007). Technical model for basic and complex emotions. In 5th IEEE International Conference on Industrial Informatics, 2:1007-1012.
[11] Damasio, A.R. (1996). Descartes Error. Papermac London.
[12] Deutsch, T.; Lang, R.; Pratl, G., Brainin, E.; and Teicher S. (2006). Applying psychoanalytic and neuro-scientific models to automation. In 2nd IET International Conference on Intelligent Environments, 1:111-118.
[13] Deutsch, T.; Gruber, A.; Lang, R.; and Velik, R. (2008). Episodic memory for autonomous agents, Conference on Human System Interactions, pp. 621-626.
[14] Deutsch, T. (2011). Human Bionically Inspired Autonomous Agents. PhD thesis, Vienna University of Technology.
[15] Dietrich, D.; and Zucker G. (2008). New approach for controlling complex processes. An introduction to the 5th generation of AI. In IEEE Conference on Human System Interactions, pp. 12-17.
[16] Dietrich, D.; Bruckner, D.; Zucker, G.; Mueller, B.; Tmej, A. (2009). Psychoanalytical model for automation and robotics. In AFRICON, pp. 1-8.
[17] Engel, A.K.; Fries, P.; and Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. Nature Reviews Neuroscience, 2.
[18] Ernst, M.O.; and Buelthoff, H.H. 2004. Merging the senses into a robust percept. TRENDS in Cognitive Sciences, 8(4):162-169.
[19] French, R.M. (1996). Review of The Engine of Reason, the Seat of the Soul. Minds and Machines, 6:416-421.
[20] Goldstein, E.B. (2002). Wahrnehmungspsychologie. Spektrum Akademischer Verlag.
[21] Golledge, H.D.R.; Hilgetag, C.C.; and Tovee, M.J. (1996). Information processing: A solution to the binding problem? Current Biology, 6(9):1092-1095.
[22] Hommel, B.; and Colzato, L.S. (2009). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. Visual Cognition, 17:120-140.
[23] Hubel, D.H.; and Wiesel, T.N. (1962). Receptive Fields, Binocular Interaction and Functional Architecture in the Cats Visual Cortex. J. Physiol., 160:106-154.
[24] Lang, R.; Bruckner, D.; Velik, R.; and Deutsch, T. (2007). Scenario Recognition in Modern Building Automation. In: Proceedings of the 7th IFAC International Conference on Fieldbuses & Networks in Industrial & Embedded Systems, pp. 305–312.
[25] Lang, R.; Bruckner, D.; Velik, R.; and Deutsch, T. (2009). Scenario Recognition in Modern Building Automation. International Journal of Intelligent Systems and Technologies, 4(1):36-44.
[26] Lorenz, B. and Barnard, E. (2007). A brief overview of Artificial Intelligence focusing on computational models of emotions. In 5th IEEE International Conference on Industrial Informatics, volume 2.
[27] Luria, A.R. (1973). The Working Brain: An Introduction in Neuropsychology. Basic Books.
[28] Palensky, B. (2008). Introducing neuro-psychoanalysis towards the design of cognitive and affective automation systems, Ph.D. dissertation, Vienna Univ. Technol., Vienna, Austria.
[29] Panksepp, J. (1998). Affective Neuroscience. Oxford University Press New York.
[30] Pongratz, M.; and Velik, R.; and Machajdik, J. (2011). Spatio-Temporal Perception Nets. Proceedings of Africon.
[31] Pratl, G. (2006). Processing and Symbolization of Ambient Sensor Data. PhD thesis, Vienna University of Technology.
[32] Qadeer, N.; Velik, R.; Zucker, G.; and Boley, H. (2009). Knowledge representation for a neuro-symbolic network in home care risk identification. In 7th IEEE International Conference on Industrial Informatics.
[33] Quian Quiroga, R.; Reddy, L.; Kreiman, G.; Koch, C.; and Fried, I. (2007). Invariant visual representation by single neurons in the human brain. Nature, 435:1102-1107.
[34] Roesener, C. (2007). Adaptive behavior arbitration for mobile service robots in building automation, Ph.D. dissertation, Vienna Univ. Technol., Vienna, Austria.

[35] Russ, G. (2003). Situation-dependent Behavior in Building Automation. PhD thesis, Vienna University of Technology.

[36] Scheutz, M.; Sloman, A.; and Logan, B. (2000). Emotional states and realistic agent behaviour. Proceedings of GameOn, pp. 81-88.

[37] Sendhoff, B.; Körner, E.; and Sporns, O. (2009). Creating Brain-Like Intelligence: From Basic Principles to Complex Intelligent Systems, chapter Creating Brain-Like Intelligence, pages 1-14. Springe.

[38] Sharkey, A.J.C. (1998). Combining Artificial Neural Nets Ensemble and Modular Multi-Net Systems. Springer-Verlag, June.

[39] Silverstein, A.; Silverstein, V.B.; and Naylor, P. (1970). Bionics: Man Copies Nature's Machines. McCall Publishing Company.

[40] Sloman, A.; and Croucher. M. (1981). Why robots will have emotions. In 7th IEEE International Joint Conference on AI.

[41] Solms, M.; and Turnbull, O. (2002). The Brain and the InnerWorld: An Introduction to the Neuroscience of Subjective Experience, Other Press, New York, NY, USA.

[42] Tamarit, C.; Dietrich, D.; Dimond, K.; and Russ, G. (2001). A definition and a model of a perceptive awareness system (PAS). In IFAC FeT, pp. 177-182.

[43] Tamarit, C. (2003). Automation System Perception: First Step towards Perceptive Awareness. PhD thesis, Vienna University of Technology.

[44] Treisman. A.M.; and Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12:97-136.

[45] Velik, R. (2007). A model for multimodal humanlike perception based on modular hierarchical symbolic information processing, knowledge integration, and learning. In 2nd IEEE International Conference on Bio-Inspired Models of Network, Information and Computing Systems, pp. 168-175.

[46] Velik, R. (2008). A Bionic Model for Human-like Machine Perception. PhD thesis, Vienna University of Technology.

[47] Velik, R. (2008). A Bionic Model for Human-like Machine Perception. Suedwestdeutscher Verlag fuer Hochschulschriften.

[48] Velik, R. 2009. Neuronal Pattern- and Object Recognition: Methods of Analysis in Time and Frequency Domain. VDM Verlag.

[49] Velik, R. (2010). Quo vadis, intelligent machine? BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 1(2):1-9.

[50] Velik, R. (2010). Towards Human-like Machine Perception 2.0. International Review on Computers and Software (IRECOS), Special Section on Advanced Artificial Networks.

[51] Velik, R. (2010). The neuro-symbolic code of perception. Journal of Cognitive Science, 11(2):161-180.

[52] Velik, R. (2010). Why Machines Cannot Feel. Minds and Machines, Springer, 20(1): 1-18.

[53] Velik, R. and Boley, H. (2010). Neurosymbolic alerting rules. IEEE Transactions on Industrial Electronics, 57(11):3661-3668.

[54] Velik, R. and Bruckner, D. (2008). Neuro-symbolic networks: Introduction to a new information processing principle. In 6th IEEE International Conference on Industrial Informatics, pp. 1042-1047.

[55] Velik, R. and Bruckner, D. (2009). A bionic approach to dynamic, multimodal scene perception and interpretation in buildings. International Journal of Intelligent Systems and Technologies, 4(1):1-9.

[56] Velik, R.; and Zucker, G. (2010). Autonomous perception and decision-making in building automation. IEEE Transactions on Industrial Electronics, 57(11):3645-3652.

[57] Velik, R.; Bruckner, D.; Lang, R.; and Deutsch, T. (2008). Emulating the perceptual system of the brain for the purpose of sensor fusion. IEEE International Conference on Human-Computer Systems Interaction, pp. 17-27.

[58] Velik, R.; Bruckner, D.; Lang, R.; and Deutsch, T. (2009). Human-Computer System Interaction: Backgrounds and Applications, chapter Emulating the Perceptual System of the Brain for the Purpose of Sensor Fusion, pp. 17-27. Springer Berlin/Heidelberg.

[59] Velik, R.; Zucker, G.; and Dietrich, D. (2011). Towards automation 2.0: A neuro-cognitive model for environment recognition, decision-making, and action execution. EURASIP Journal on Embedded Systems, 2011(11).

[60] Velik, R. (2012). From simple receptors to complex multimodal percepts: A first global picture on the mechanisms involved in perceptual binding. Frontiers in Cognitive Science, 3:1-13, 2012.

[61] Velik, R. (2012). AI Reloaded: Objectives, Potentials, and Challenges of the Novel Field of Brain-Like Artificial Intelligence. BRAIN Broad Research in Artificial Intelligence and Neuroscience, 3(3): 25–54.

[62] Velik, R. (2013). Cognitive Architectures as Building Energy Management System for Future Renewable Energy Scenarios; A Work in Progress Report, International Journal of Science and Engineering Investigations, 2(17):67-72.

[63] Velik, R. (submitted 2013). A Brain-Inspired Multimodal Data Mining Approach for Human Activity Recognition in Elderly Homes. 21 pages.

[64] Velik, R. (in preparation 2013). Cognitive Decision Agent Architectures for Individualized, Goal-Specific Energy Management and Trading in Prosumer Buildings and Neighbourhoods.

[65] Wiener, N. (2007). Cybernetics or Control and Communication in the Animal and the Machine. Kessinger Publishing, LLC.

[66] Wolfe J.M.; and Cave, K.R. (1999). The psychophysical evidence for a binding problem. Neuron, 24:11-17.